

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 5, 2004, 05:30:38 ; Search time 14364 Seconds
(without alignments)
11516.247 Million cell updates/sec

Title: US-09-679-643-1
Perfect score: 3498
Sequence: 1 ggtaccataataacacatc.....gaagagtataagaagaagctt 3498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3498	100.0	3498	6	AR096793 Sequence
2	3498	100.0	3498	6	AR125115 Sequence
3	3487	99.7	3499	1	SACFG
4	3416.2	97.7	37669	1	AC074317
5	3399.6	97.2	4709	6	AR354469 Sequence
6	3277	93.7	110000	1	Continuation (9 of
7	3245	92.8	290150	1	AP004824 Staphyloc
8	2960.6	84.6	348477	1	AP003360 Staphyloc
9	2853.4	81.6	299050	1	AP003131 Staphyloc
10	2797.4	80.0	2799	6	AX583664 Sequence
11	2746.6	78.5	2781	6	AX619958 Sequence
12	2686.8	76.8	110000	1	Continuation (9 of
13	1776.4	50.8	2792	6	AX343875 Sequence
14	1776.4	50.8	2792	6	AX768506 Sequence
15	917.8	26.2	2721	1	AF404823 Staphyloc
16	619.2	17.7	5202	1	AF245041 Staphyloc
17	619.2	17.7	5406	6	BD245129 Polypepti
18	619.2	17.7	5406	6	AR410982 Sequence
19	619.2	17.7	5406	6	AR482275 Sequence

C	20	592.8	16.9	90802	1	AE016752 Staphyloc
	21	589.6	16.9	4627	1	AF269905 Staphyloc
	22	589.6	16.9	4627	6	AR485859 Sequence
	23	589.6	16.9	4627	6	AX145223 Sequence
	24	584.8	16.7	1485	6	AR483599 Sequence
	25	584.8	16.7	1485	6	AX141313 Sequence
	26	583.4	16.7	4656	1	AY048593 Staphyloc
C	27	574.4	16.4	117662	1	AP004831 Staphyloc
	28	568.6	16.3	2969	6	AR453524 Sequence
	29	568.6	16.3	2969	6	BD140510 Fibrinoge
	30	568.6	16.3	2971	1	SAA224764
	31	568.4	16.2	2133	1	AY510088 Staphyloc
	32	568	16.2	2739	6	AX583634 Sequence
	33	562	16.1	269050	1	AL935255 Lactobaci
	34	558.2	16.0	1059	6	AX821821 Sequence
C	35	555.6	15.9	110000	1	Continuation (27 o
C	36	541.8	15.5	99802	1	Continuation (28 o
	37	525.4	15.0	2151	6	AX583629 Sequence
	38	510	14.6	2631	6	AX620920 Sequence
C	39	510	14.6	3007	6	AX768448 Sequence
	40	510	14.6	11050	6	AR353978 Sequence
C	41	510	14.6	49898	1	AC027136 Staphyloc
C	42	502.6	14.4	2582	6	AX768449 Sequence
C	43	502.6	14.4	2582	6	AX768519 Sequence
C	44	502	14.4	109528	1	AP003366 Staphyloc
C	45	502	14.4	117016	1	AP003138 Staphyloc

ALIGNMENTS

RESULT 1
AR096793
LOCUS AR096793 3498 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 1 from patent US 6008341.
ACCESSION AR096793
VERSION AR096793.1 GI:10025907
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3498)
AUTHORS Foster,T.James. and McDevitt,D.Leo.
TITLE S. aureus fibrinogen binding protein gene
JOURNAL Patent: US 6008341-A 1 28-DEC-1999;
FEATURES
Location/Qualifiers
1..3498
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	100.0%	Score	3498	DB	6	Length	3498
Best Local Similarity	100.0%	Pred. No.	0				
Matches	3498	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	GGTACCATTAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTCATTA	60				
Db	1	GGTACCATTAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTCATTA	60				
QY	61	GGTAGAGTTTCATATTAAATAATAAATAAATGTTTGCAATCAATCGTACGTTGCTGTTTGA	120				
Db	61	GGTAGAGTTTCATATTAAATAATAAATAAATGTTTGCAATCAATCGTACGTTGCTGTTTGA	120				
QY	121	ATTCCTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA	180				
Db	121	ATTCCTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA	180				
QY	181	TATCGATACAAATTAATTCGCTATATGCAATTTTAGTGTATTAATTCATTAACAGAGATT	240				
Db	181	TATCGATACAAATTAATTCGCTATATGCAATTTTAGTGTATTAATTCATTAACAGAGATT	240				
QY	241	AAATATATCTTAAAGGGTATATAGTTATATATAAATGACCTTTTAAAGAGGGATATAA	300				

applicants

Db 241 AAATAATCTTAAAGGCTATATAGTTAATAATAAATGACTTTTAAAGAGGGAATAAA 300
QY
301 ATGAAATGAAGAAAGAAAGAAACACGCAATTCGGAAAGAAATCGATTGGCGTTCAC 360
Db 301 ATGAAATGAAGAAAGAAAGAAACACGCAATTCGGAAAGAAATCGATTGGCGTTCAC 360
QY 361 GTGCTGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 420
Db 361 GTGCTGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAGT 420
QY 421 GAAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATCAAGT 480
Db 421 GAAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATCAAGT 480
QY 481 AGCGTTAGTGTCACTAAACAGACGACACAAACGAGTGTAGTACTATAAACATCGTCA 540
Db 481 AGCGTTAGTGTCACTAAACAGACGACACAAACGAGTGTAGTACTATAAACATCGTCA 540
QY 541 AACACTAATATGCGGAAACGAGTGTGGCGCAAAATCCAGCACACAGGAAACGACACAA 600
Db 541 AACACTAATATGCGGAAACGAGTGTGGCGCAAAATCCAGCACACAGGAAACGACACAA 600
QY 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGTGTGAGCTACTACTACG 660
Db 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGTGTGAGCTACTACTACG 660
QY 661 ACACGATCAAGCTAATACACCGGCAACACTCAATCAAGCAATACAAATGCGGAGAA 720
Db 661 ACACGATCAAGCTAATACACCGGCAACACTCAATCAAGCAATACAAATGCGGAGAA 720
QY 721 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA 780
Db 721 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA 780
QY 781 AATTCACTCAAAATCTTACAAATGCGGAAATGTTTCAACAGCGAAGTACTTCAACT 840
Db 781 AATTCACTCAAAATCTTACAAATGCGGAAATGTTTCAACAGCGAAGTACTTCAACT 840
QY 841 GAAGCAACACTTCAACAAATGAATCAGTCCAGAGTACAGATGCAAGTAATAAGAT 900
Db 841 GAAGCAACACTTCAACAAATGAATCAGTCCAGAGTACAGATGCAAGTAATAAGAT 900
QY 901 GTAGTTAATCAAGCGGTTAATACAAAGTGGCTAGAAATGAGAGCATTTAGTTAGCGCA 960
Db 901 GTAGTTAATCAAGCGGTTAATACAAAGTGGCTAGAAATGAGAGCATTTAGTTAGCGCA 960
QY 961 GTAGCTGAGATGCAACGCGGCTGGGACAGATTAATGAGATTCAGTTCAGCAATGTGACA 1020
Db 961 GTAGCTGAGATGCAACGCGGCTGGGACAGATTAATGAGATTCAGTTCAGCAATGTGACA 1020
QY 1021 GTTGTATTGACTCTGTGACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAAT 1080
Db 1021 GTTGTATTGACTCTGTGACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAAT 1080
QY 1081 TATGTTTTCAGTGCCTAAATCTGCTGTTAAAGGTGACACATTCAAAATAACTGTACCT 1140
Db 1081 TATGTTTTCAGTGCCTAAATCTGCTGTTAAAGGTGACACATTCAAAATAACTGTACCT 1140
QY 1141 AAAGAAATTAACCTTAAATGGTGTAACTTCACTGCTAAAGTGCCACCAATTTATGGCTGGA 1200
Db 1141 AAAGAAATTAACCTTAAATGGTGTAACTTCACTGCTAAAGTGCCACCAATTTATGGCTGGA 1200
QY 1201 GATCAAGTATTGGCAAAATGGTGTAAATCGATAGTGTGTTATTTATACATTTACA 1260
Db 1201 GATCAAGTATTGGCAAAATGGTGTAAATCGATAGTGTGTTATTTATACATTTACA 1260
QY 1261 GACTATGTAAATTAACAGATGATAAGCAACTTTGACCAATGCCCGCTTATATTGAC 1320
Db 1261 GACTATGTAAATTAACAGATGATAAGCAACTTTGACCAATGCCCGCTTATATTGAC 1320
QY 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1380
Db 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1380

QY 1381 GCAACAAAAACAGTATTAGTAGATTATGAAAAATATAGTAAAGTTTATAAATTATCATT 1440
Db 1381 GCAACAAAAACAGTATTAGTAGATTATGAAAAATATAGTAAAGTTTATAAATTATCATT 1440
QY 1441 AAAGGTACAATTCGACCAATTCGATAAAACAAATTAATACGTATCGTCAGACAAATTTATGTC 1500
Db 1441 AAAGGTACAATTCGACCAATTCGATAAAACAAATTAATACGTATCGTCAGACAAATTTATGTC 1500
QY 1501 AATCCAGTGGAGATACGTTATTGGCCGGTTTAAACAGTAAATTTAAACACCAAAATACG 1560
Db 1501 AATCCAGTGGAGATACGTTATTGGCCGGTTTAAACAGTAAATTTAAACACCAAAATACG 1560
QY 1561 GATAGTAAATGATTAATAGATCAGCAAAATAACAAGTATTAAAGTATATAAAGTAGATAAT 1620
Db 1561 GATAGTAAATGATTAATAGATCAGCAAAATAACAAGTATTAAAGTATATAAAGTAGATAAT 1620
QY 1621 GCAGCTGATTACTGAAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACTAAT 1680
Db 1621 GCAGCTGATTACTGAAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACTAAT 1680
QY 1681 AGTGTGAATATTACATTCACAAATCCAAATCAATATAAAGTAGAGTTTAAATACCCCTGAT 1740
Db 1681 AGTGTGAATATTACATTCACAAATCCAAATCAATATAAAGTAGAGTTTAAATACCCCTGAT 1740
QY 1741 GATCAAAATTACAAACACCGGTATATAGTAGTTTAAATGTCATATTTGATCCGAAATAGCAAA 1800
Db 1741 GATCAAAATTACAAACACCGGTATATAGTAGTTTAAATGTCATATTTGATCCGAAATAGCAAA 1800
QY 1801 GGTGATTAGCTTTACGTTCACTTTATATGGGTATTAACTCGAATATAATTTGGCGTCT 1860
Db 1801 GGTGATTAGCTTTACGTTCACTTTATATGGGTATTAACTCGAATATAATTTGGCGTCT 1860
QY 1861 ATGTCATGGGACACAGTAGCAATTTAAACCGATCAGGTTCTGTGTGACGGTATCGAT 1920
Db 1861 ATGTCATGGGACACAGTAGCAATTTAAACCGATCAGGTTCTGTGTGACGGTATCGAT 1920
QY 1921 AAACAGTGTCTCTGAAACCTGATGAGCTTGGTGAATTTGAACCAATTTCCAGAGAT 1980
Db 1921 AAACAGTGTCTCTGAAACCTGATGAGCTTGGTGAATTTGAACCAATTTCCAGAGAT 1980
QY 1981 TCAGATTCTGACCCAGGTTAGATTCTGGGACGATTTCTAATTCAGTAGCGGTTTCAGAT 2040
Db 1981 TCAGATTCTGACCCAGGTTAGATTCTGGGACGATTTCTAATTCAGTAGCGGTTTCAGAT 2040
QY 2041 TCGGTTAGTGAATCTCATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 2100
Db 2041 TCGGTTAGTGAATCTCATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA 2100
QY 2101 AGTGATTTCAGACTCAGCGAGTGAATTCAGATTTCAGCAAGCGATTCGACTCAGCGAGCGAT 2160
Db 2101 AGTGATTTCAGACTCAGCGAGTGAATTCAGATTTCAGCAAGCGATTCGACTCAGCGAGCGAT 2160
QY 2161 TCCGACTCAGCAATGACTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGAT 2220
Db 2161 TCCGACTCAGCAATGACTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGAT 2220
QY 2221 TCCGACAGTGAATTCAGATTTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT 2280
Db 2221 TCCGACAGTGAATTCAGATTTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT 2280
QY 2281 AGCGATTTCAGATTTCAGATAGCGATTTTCAGATTTCGAGTTCGACTCAGACAGCGAT 2340
Db 2281 AGCGATTTCAGATTTCAGATAGCGATTTTCAGATTTCGAGTTCGACTCAGACAGCGAT 2340
QY 2341 TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTTCGAGTTCGACTCAGACAG 2400
Db 2341 TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTTCGAGTTCGACTCAGACAG 2400
QY 2401 TCAGATAGCGATTTCGACTCAGATAGCGACTCAGATTTCAGACAGCGATTCAGATTCAGAT 2460
Db 2401 TCAGATAGCGATTTCGACTCAGATAGCGACTCAGATTTCAGACAGCGATTCAGATTCAGAT 2460

Db ||||| 721 TTAGTGAATCAACAAGTAATGAACACGCTTTAAATGATACTAATACAGTATCATCTGTA 780
Qy ||||| 781 AATTCACTCAAAATTTCTCAAAATCGGAAATATGTTTCAACAAACGCAAGATATCTCAACT 840
Db ||||| 781 AATTCACTCAAAATTTCTCAAAATCGGAAATATGTTTCAACAAACGCAAGATATCTCAACT 840
Qy ||||| 841 GAAGCAACACTTCAAAATGAATCAGCTCCACAGAGTACAGATGCAAGTCAATTAAGAT 900
Db ||||| 841 GAAGCAACACTTCAAAATGAATCAGCTCCACAGAGTACAGATGCAAGTCAATTAAGAT 900
Qy ||||| 901 GTAGTTAATCAAGCGGTTAATACAAGTGGCTTGAATGAGAGCATTTAGTTTACGGCA 960
Db ||||| 901 GTAGTTAATCAAGCGGTTAATACAAGTGGCTTGAATGAGAGCATTTAGTTTACGGCA 960
Qy ||||| 961 GTAGCTCAGATGACCGGAGCTGGCACAGATATTACGAATCATGTTGACGAATGTGACA 1020
Db ||||| 961 GTAGCTCAGATGACCGGAGCTGGCACAGATATTACGAATCATGTTGACGAATGTGACA 1020
Qy ||||| 1021 GTTGGTATTGACTCTGTGACGACTGTATCCGCAACAGCAGGTTATGTCAAACTGAAT 1080
Db ||||| 1021 GTTGGTATTGACTCTGTGACGACTGTATCCGCAACAGCAGGTTATGTCAAACTGAAT 1080
Qy ||||| 1081 TATGTTTTTTCAGTGCCTTAATTTCTGCTTAAAGTGACACATTCAAAATAACTGTACCT 1140
Db ||||| 1081 TATGTTTTTTCAGTGCCTTAATTTCTGCTTAAAGTGACACATTCAAAATAACTGTACCT 1140
Qy ||||| 1141 AAAGAAATTAATGTTAAATGGTAACTTCAACTGCTAAAGTGACACATTCAAAATAACTGTGGA 1200
Db ||||| 1141 AAAGAAATTAATGTTAAATGGTAACTTCAACTGCTAAAGTGACACATTCAAAATAACTGTGGA 1200
Qy ||||| 1201 GATCAAGTATTGGCAATGGTAAATCGATAGTGAATGTTAATGTTAATACATTTACA 1260
Db ||||| 1201 GATCAAGTATTGGCAATGGTAAATCGATAGTGAATGTTAATGTTAATACATTTACA 1260
Qy ||||| 1261 GACTATGTAATATCTAAAGATGATGTAAGAACAACTTTGACCATGCCCGCTTATATTGAC 1320
Db ||||| 1261 GACTATGTAATATCTAAAGATGATGTAAGAACAACTTTGACCATGCCCGCTTATATTGAC 1320
Qy ||||| 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAA 1380
Db ||||| 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAA 1380
Qy ||||| 1381 GCAACAAAACAGTATTAGTATGATGTAAGAAATATGTTTAACTTATCTATT 1440
Db ||||| 1381 GCAACAAAACAGTATTAGTATGATGTAAGAAATATGTTTAACTTATCTATT 1440
Qy ||||| 1441 AAAGTCAATTTGACCAATCGATAAAACAAATAATACGTATCGTCAGACAAATTTATGTC 1500
Db ||||| 1441 AAAGTCAATTTGACCAATCGATAAAACAAATAATACGTATCGTCAGACAAATTTATGTC 1500
Qy ||||| 1501 AATCCAAGTGAGATAACGTTATTTCGGCGGTTTTTAAACAGGTAATTTAAACCAATACG 1560
Db ||||| 1501 AATCCAAGTGAGATAACGTTATTTCGGCGGTTTTTAAACAGGTAATTTAAACCAATACG 1560
Qy ||||| 1561 GATAGTAATGATTAATAGATCAGCAATACAGTATTAAAGTATATAAGTATAGATAAT 1620
Db ||||| 1561 GATAGTAATGATTAATAGATCAGCAATACAGTATTAAAGTATATAAGTATAGATAAT 1620
Qy ||||| 1621 GCAGCTGATTTATCTGAAAGTTACTTTTGTGAATCCAGAAACTTTGAGGATGTCCTAAT 1680
Db ||||| 1621 GCAGCTGATTTATCTGAAAGTTACTTTTGTGAATCCAGAAACTTTGAGGATGTCCTAAT 1680
Qy ||||| 1681 AGTGTGAATATTAATCCCAATCCAAATCAATTAAGTATAGATTTAATACGCTGAT 1740
Db ||||| 1681 AGTGTGAATATTAATCCCAATCCAAATCAATTAAGTATAGATTTAATACGCTGAT 1740
Qy ||||| 1741 GATCAAAATTAACACCGTATATAGTATGTTTAAATGCTCATATTCATCCGAATAGCAAA 1800
Db ||||| 1741 GATCAAAATTAACACCGTATATAGTATGTTTAAATGCTCATATTCATCCGAATAGCAAA 1800
Qy ||||| 1801 GGTGATTTAGCTTTAGCTTCAACTTTTATATGGGTATTAATCTGCAATATAATTTTGGCGCTCT 1860
Db ||||| 1801 GGTGATTTAGCTTTAGCTTCAACTTTTATATGGGTATTAATCTGCAATATAATTTTGGCGCTCT 1860
Qy ||||| 1861 ATGTCAATGGACAAACAGTAGCATTTAAATAACCGATCAGGTTCTGTGTGACGATATCGAT 1920
Db ||||| 1861 ATGTCAATGGACAAACAGTAGCATTTAAATAACCGATCAGGTTCTGTGTGACGATATCGAT 1920
Qy ||||| 1921 AAACCAAGTTCTTCTCTGAAACAACTGATGAGCTGGTGAATTTGAACCAATTTCCAGAGAT 1980
Db ||||| 1921 AAACCAAGTTCTTCTCTGAAACAACTGATGAGCTGGTGAATTTGAACCAATTTCCAGAGAT 1980
Qy ||||| 1981 TCAGATTTCTACCCAGGTTTCAAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTCAAT 2040
Db ||||| 1981 TCAGATTTCTACCCAGGTTTCAAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTCAAT 2040
Qy ||||| 2041 TCGGTTAGTGTATTTACATCAGATAGTGGTTTCAAGTTTCAAGAGGATTTCCAGATTCAGCA 2100
Db ||||| 2041 TCGGTTAGTGTATTTACATCAGATAGTGGTTTCAAGTTTCAAGAGGATTTCCAGATTCAGCA 2100
Qy ||||| 2101 AGTGATTCAGACTCAGCGAGTGTATTCAGATTTCAAGAAAGCGATTTCCGACTCAGCGAGCGAT 2160
Db ||||| 2101 AGTGATTCAGACTCAGCGAGTGTATTCAGATTTCAAGAAAGCGATTTCCGACTCAGCGAGCGAT 2160
Qy ||||| 2161 TCCGACTCAGACAAATGATTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGTATTCAGAT 2220
Db ||||| 2161 TCCGACTCAGACAAATGATTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGTATTCAGAT 2220
Qy ||||| 2221 TCCGACTCAGACAAATGATTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGTATTCAGAT 2280
Db ||||| 2221 TCCGACTCAGACAAATGATTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGTATTCAGAT 2280
Qy ||||| 2281 AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGCGAT 2340
Db ||||| 2281 AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGCGAT 2340
Qy ||||| 2341 TCTGACTCCGACAGTGTATTCGACTCAGACAGCGATTTCCGACAGTGTATTCGACTCAGAC 2400
Db ||||| 2341 TCTGACTCCGACAGTGTATTCGACTCAGACAGCGATTTCCGACAGTGTATTCGACTCAGAC 2400
Qy ||||| 2401 TCAGATAGCGATTTCCGACTCAGATAGCGATTTCCGACAGCGATTTCCGACTCAGATTCAGAC 2460
Db ||||| 2401 TCAGATAGCGATTTCCGACTCAGATAGCGATTTCCGACAGCGATTTCCGACTCAGATTCAGAC 2460
Qy ||||| 2461 AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGTGTATTCGACT 2520
Db ||||| 2461 AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGTGTATTCGACT 2520
Qy ||||| 2521 TCGGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGTGTATTCGACTCAGAC 2580
Db ||||| 2521 TCGGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGTGTATTCGACTCAGAC 2580
Qy ||||| 2581 TCAGACAGTGTATTCGACTCAGACAGTGTATTCGACTCAGATAGTGTATTCGACTCAGAC 2640
Db ||||| 2581 TCAGACAGTGTATTCGACTCAGACAGTGTATTCGACTCAGATAGTGTATTCGACTCAGAC 2640
Qy ||||| 2641 AGTGATTCGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGATAGTGTATTCGACT 2700
Db ||||| 2641 AGTGATTCGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGATAGTGTATTCGACT 2700
Qy ||||| 2701 TCGGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGCGATTTCCGACTCAGAT 2760
Db ||||| 2701 TCGGATTCAGATAGCGATTTCCGACAGTGTATTCGACTCAGACAGCGATTTCCGACTCAGAT 2760
Qy ||||| 2761 TCAGACAGCGATTCAGACAGTGTATTCGACTCAGATAGTGTATTCGACTCAGCGAGTGTAT 2820
Db ||||| 2761 TCAGACAGCGATTCAGACAGTGTATTCGACTCAGATAGTGTATTCGACTCAGCGAGTGTAT 2820
Qy ||||| 2821 TCAGACTCAGGTAAGTGTATTCGATTTCAAGTGTATTCGACTCAGAAAGTGTATTCAGAT 2880
Db ||||| 2821 TCAGACTCAGGTAAGTGTATTCGATTTCAAGTGTATTCGACTCAGAAAGTGTATTCAGAT 2880
Qy ||||| 2881 AGCGATTCGAGTGTATTCGACTCAGAAAGTGTATTCGACTCAGAAAGTGTATTCGACTCAGAA 2940
Db ||||| 2881 AGCGATTCGAGTGTATTCGACTCAGAAAGTGTATTCGACTCAGAAAGTGTATTCGACTCAGAA 2940


```
QY 2941 ACTAATGCTCTCTAATAAATAGAGCTAAAGATAGTAAGAACCAATTTACCAGATACAGGT 3000
|
|
|
Db 2941 ACTAATGCTCTCTAATAAATAGAGCTAAAGATAGTAAGAACCAATTTACCAGATACAGGT 3000
|
|
|
QY 3001 TCTGAAGATGAAGCAAAATACGTCACCTAATTTGGGATTAATAGCATCAATFAGGTTCAATTA 3060
|
|
|
Db 3001 TCTGAAGATGAAGCAAAATACGTCACCTAATTTGGGATTAATAGCATCAATFAGGTTCAATTA 3060
|
|
|
QY 3061 CTACTTTTCAGAGAGAAAAGAAAATAAGATAAGAAAATAAGTAATATGATTAATTAAT 3120
|
|
|
Db 3061 CTACTTTTCAGAGAGAAAAGAAAATAAGATAAGAAAATAAGTAATATGATTAATTAAT 3120
|
|
|
QY 3121 TAATCATATGATTCAATGAAGAACCCCTTAAAGAGTGCTTCTTTTACTTTGGATTTTCCA 3180
|
|
|
Db 3121 TAATCATATGATTCAATGAAGAACCCCTTAAAGAGTGCTTCTTTTACTTTGGATTTTCCA 3180
|
|
|
QY 3181 AATATATTTGTTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240
|
|
|
Db 3181 AATATATTTGTTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240
|
|
|
QY 3241 TAGATGTTATTAATTTGCTTGGCGAAGAAATAGGTTGAAGTTAGTTTAAATAGG 3300
|
|
|
Db 3241 TAGATGTTATTAATTTGCTTGGCGAAGAAATAGGTTGAAGTTAGTTTAAATAGG 3300
|
|
|
QY 3301 GAAATTAAGGAGAAAATACAGTTGAAAATAAATTAATTTGCTAGTTTATCATTTGGAGCATT 3360
|
|
|
Db 3301 GAAATTAAGGAGAAAATACAGTTGAAAATAAATTTGCTAGTTTATCATTTGGAGCATT 3360
|
|
|
QY 3361 ATGTGTATCACAAATTTGGGAAAGTAATCGTGGAGTGCAAGTGCTTTCTGGGAGAGAA 3420
|
|
|
Db 3361 ATGTGTATCACAAATTTGGGAAAGTAATCGTGGAGTGCAAGTGCTTTCTGGGAGAGAA 3420
|
|
|
QY 3421 TCATATGATCTGAGCTCGTTGAACTGACTATATAAATAAATTAATTAATTAATTAATTAAT 3480
|
|
|
Db 3421 TCATATGATCTGAGCTCGTTGAACTGACTATATAAATAAATTAATTAATTAATTAATTAAT 3480
|
|
|
QY 3481 AGAGTATAGAAAAGCTT 3498
|
|
|
Db 3481 AGAGTATAGAAAAGCTT 3498
|
|
|
RESULT 3
SACFG 3499 bp DNA linear BCT 25-OCT-1994
LOCUS S.aureus gene for clumping factor.
DEFINITION 21852
VERSION 21852.1 Gi:397525
KEYWORDS clumping factor; fibrinogen receptor.
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1 (bases 1 to 3499)
AUTHORS McDevitt,D., Francois,P., Vaudaux,P. and Foster,T.J.
TITLE Molecular characterization of the clumping factor (fibrinogen
receptor) of Staphylococcus aureus
JOURNAL Mol. Microbiol. 11 (2), 237-248 (1994)
MEDLINE 94224142
PUBMED 8170386
REFERENCE 2 (bases 1 to 3499)
AUTHORS McDevitt,D.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1992) McDevitt D., Trinity College, Microbiology,
Dublin, Ireland, 2
FEATURES Location/Qualifiers
source 1..3499
/organism="Staphylococcus aureus"
/mol_type="genomic DNA"
/strain="Newman"
/db_xref="taxon:1280"
302..3103
/function="Fibrinogen receptor"
/citation=[1]
```

```
/codon_start=1
/translation="NMKKKKEHARRKSGIVASVUVGLIGRLGLSSKEADASENSV
TUSDASNSKSDSSSAAPKTDITNSVSDTKTNSNTNNNGTSSVAQNPAQOETTQSS
SNATTEETPTVTEATTTTTNQANTPATQSSNTNABEELNQTNSNTFNNTTNTSVS
NSPQSTNAEASTTQDTTEATPSNNEGAPQSTDAKNDVNVQAVNTAPRMRATSL
AAVADAAPAGTDITNLTNTVTGIDSDGTTVPHOAGYVKLYKLVGSPVNSAVGDDTFK
ITVPEKLNGLNGVTSTAKVPIAGDOVLANGVIDSDGNVLYTFDVVNTKDDV/KALIT
MAYIDPENVKKTGNTLATGISTGTANTKTLVLDYKGYKFNLSIKGIDQDKZNN
TRQTIYNPSGDNVIAPIATGLNKENTSNALIDQNTSIKIVKVDNAADLSESVFV
NFTEDVTNNSIITFPNPKYKFEPTDQITTPYIVVNGHIDPNKGLALAST
LYGANSNITWRSMWMDNEVAFNGSGDGIKDPVPEQEPGEPIPEOSDSDPG
SDGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSE
SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSE
SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSE
ASIGSLLLFRRKKENKDKK"
ORIGIN
Query Match 99.7%; Score 3487; DB 1; Length 3499;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3498; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGTACCATAAATTACACATCTGCTTTTGAAAATAATATGATTTCAAGCTAGGATACATTA 60
|
|
|
Db 1 GGTACCATAAATTACACATCTGCTTTTGAAAATAATATGATTTCAAGCTAGGATACATTA 60
|
|
|
QY 61 GGTAGAGTTTCATATTAATAAATAAATGTTTGCAATCAAAATCGTACGTTGTCGTTGTA 120
|
|
|
Db 61 GGTAGAGTTTCATATTAATAAATAAATGTTTGCAATCAAAATCGTACGTTGTCGTTGTA 120
|
|
|
QY 121 ATCTTTAAATAGCAATAAATAAATGTTTGTTAGTAAAGTATATTGTTGGATAATAAAA 180
|
|
|
Db 121 ATCTTTAAATAGCAATAAATAAATGTTTGTTAGTAAAGTATATTGTTGGATAATAAAA 180
|
|
|
QY 181 TATCGATACAAATTAATTCCTATTAATGCAATTTTACTGTATATTCATTACAGAGATT 240
|
|
|
Db 181 TATCGATACAAATTAATTCCTATTAATGCAATTTTACTGTATATTCATTACAGAGATT 240
|
|
|
QY 241 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTAAAAAGGGGAATAA 299
|
|
|
Db 241 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTAAAAAGGGGAATAA 300
|
|
|
QY 300 AATGAATATGAAGAAAAGAAAACACGCAATTCGGAATAAATCGATTGGCGTGCTTC 359
|
|
|
Db 301 AATGAATATGAAGAAAAGAAAACACGCAATTCGGAATAAATCGATTGGCGTGCTTC 360
|
|
|
QY 360 AGTGCTTTGAGGTAGCTTAATCGTTTTCGACTACTCAGCAGTAAAGAGCAGATCGCA 419
|
|
|
Db 361 AGTGCTTTGAGGTAGCTTAATCGTTTTCGACTACTCAGCAGTAAAGAGCAGATCGCA 420
|
|
|
QY 420 TGAATAATAGTTTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTCAAG 479
|
|
|
Db 421 TGAATAATAGTTTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTCAAG 480
|
|
|
QY 480 TAGCGTTAGTGTGCACCTTAAACAGACGACACAAACGTTGAGTACTATAAACATCGTC 539
|
|
|
Db 481 TAGCGTTAGTGTGCACCTTAAACAGACGACACAAACGTTGAGTACTATAAACATCGTC 540
|
|
|
QY 540 AAACACTAATATGCGGAACAGGTGTGCGCGAAAATCCAGCACACACAGAAACGACACA 599
|
|
|
Db 541 AAACACTAATATGCGGAACAGGTGTGCGCGAAAATCCAGCACACACAGAAACGACACA 600
|
|
|
QY 600 ATCATCATCAACAAATGCAACTACCGAAGAAACGCGGTAACTGTTGAGGTACTACTACTAC 659
|
|
|
```

Db 601 ATCATCATCAACAAATGCAATACGGAAGAAACGCCGGTAACTGGTGAAGCTACTACTAC 660
Qy 660 GACAAACGAATCAAGCTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 719
Db 661 GACAAACGAATCAAGCTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 720
Qy 720 ATTAGTGAATCAAAACAAAGTAATGAACGACATTTTAAATGATACTAATACAGTATCATCTGT 779
Db 721 ATTAGTGAATCAAAACAAAGTAATGAACGACATTTTAAATGATACTAATACAGTATCATCTGT 780
Qy 780 AAATTCACCTCAAAATTTCTCAAAATGCGGAAATATGTTTCAACAAACGCAAGATCTTCAAC 839
Db 781 AAATTCACCTCAAAATTTCTCAAAATGCGGAAATATGTTTCAACAAACGCAAGATCTTCAAC 840
Qy 840 TGAAGCAACCTTTCAAAACAAATGAATCAGCTCCACAGATACAGATGCAAGTAATAAGA 899
Db 841 TGAAGCAACCTTTCAAAACAAATGAATCAGCTCCACAGATACAGATGCAAGTAATAAGA 900
Qy 900 TGTAGTTAATCAAGCGGTTAATACAAGTGGCGCTAGAAATGAGACATTTAGTTAGCGGC 959
Db 901 TGTAGTTAATCAAGCGGTTAATACAAGTGGCGCTAGAAATGAGACATTTAGTTAGCGGC 960
Qy 960 AGTAGCTGCAGATGCAACCGGAGCTGGCAAGATATACGAATCAGTTGACGAATGTGAC 1019
Db 961 AGTAGCTGCAGATGCAACCGGAGCTGGCAAGATATACGAATCAGTTGACGAATGTGAC 1020
Qy 1020 AGTTGGTATTCAGCTCTGTGACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAACTGAA 1079
Db 1021 AGTTGGTATTCAGCTCTGTGACGACTGTGTATCCGCAACCAAGCAGGTTATGTCAAACTGAA 1080
Qy 1080 TTATGGTTTTTCAGTGCCCTAATTCGCTGTTTAAAGGTGACACATTCAAAATFAACTGTACC 1139
Db 1081 TTATGGTTTTTCAGTGCCCTAATTCGCTGTTTAAAGGTGACACATTCAAAATFAACTGTACC 1140
Qy 1140 TAAAGAAATTAACCTTAATAGTGTTAACTTCACTGCTTAAGTGCCACCAATATGGCTGG 1199
Db 1141 TAAAGAAATTAACCTTAATAGTGTTAACTTCACTGCTTAAGTGCCACCAATATGGCTGG 1200
Qy 1200 AGATCAAGTATGGCAAAATGGTAAATCGATAGTGAATGATGTTATTTATACATTTAC 1259
Db 1201 AGATCAAGTATGGCAAAATGGTAAATCGATAGTGAATGATGTTATTTATACATTTAC 1260
Qy 1260 AGACTATGTAATACTAAAGATGATGTAAGAAAGCAACTTTGACCCGCTTATATTGA 1319
Db 1261 AGACTATGTAATACTAAAGATGATGTAAGAAAGCAACTTTGACCCGCTTATATTGA 1320
Qy 1320 CCCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGGCATAGGTAGTACAAAC 1379
Db 1321 CCCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGGCATAGGTAGTACAAAC 1380
Qy 1380 AGCAAAACAAACAGTATTAGTAGATTATGAAATATGTTAAGTTTATATACTTATCTAT 1439
Db 1381 AGCAAAACAAACAGTATTAGTAGATTATGAAATATGTTAAGTTTATATACTTATCTAT 1440
Qy 1440 TAAAGGTACAAATTCGAAATTCGATPAAACAAATTAATACGTTATCGTCAGACAAATTTATGT 1499
Db 1441 TAAAGGTACAAATTCGAAATTCGATPAAACAAATTAATACGTTATCGTCAGACAAATTTATGT 1500
Qy 1500 CAATCCAGTGGAGATACGTTATTTGCGCCGGTTTAAACAGGTAATTTAAACCAAAATAC 1559
Db 1501 CAATCCAGTGGAGATACGTTATTTGCGCCGGTTTAAACAGGTAATTTAAACCAAAATAC 1560
Qy 1560 GGATAGTAATGCAATTAATAGATACGAAATATCAAGTATTAAAGTATATAAGTAGATAA 1619
Db 1561 GGATAGTAATGCAATTAATAGATACGAAATATCAAGTATTAAAGTATATAAGTAGATAA 1620
Qy 1620 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTTGAGGATGTCACTPAA 1679
Db 1621 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTTGAGGATGTCACTPAA 1680
Qy 1680 TAGTGTGAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATCCGCTGA 1739
Db 1681 TAGTGTGAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATCCGCTGA 1740

Qy 1740 TGATCAAAATTAACAAACCGTATATAGTAGTTGTTTAAATGGTCATATTTGATCCGAATAGCAA 1799
Db 1741 TGATCAAAATTAACAAACCGTATATAGTAGTTGTTTAAATGGTCATATTTGATCCGAATAGCAA 1800
Qy 1800 AGGTGATTTAGCTTTACGTTCAACTTTATATGGGTAAATCTCGAATAATAATTTGGCGCTC 1859
Db 1801 AGGTGATTTAGCTTTACGTTCAACTTTATATGGGTAAATCTCGAATAATAATTTGGCGCTC 1860
Qy 1860 TATGTCATGGGACAAACGAAGTAGCATTTAATAACGGATCAGGTTCTGGTACGGTATCGA 1919
Db 1861 TATGTCATGGGACAAACGAAGTAGCATTTAATAACGGATCAGGTTCTGGTACGGTATCGA 1920
Qy 1920 TAAACCAAGTTGTTCTCTGAAACAAACCTGATGAGCTGGTGAATTTGAACCAATTCAGAGGA 1979
Db 1921 TAAACCAAGTTGTTCTCTGAAACAAACCTGATGAGCTGGTGAATTTGAACCAATTCAGAGGA 1980
Qy 1980 TTGAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATACGGGTTTCA 2039
Db 1981 TTGAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATACGGGTTTCA 2040
Qy 2040 TTGCGGTAGTAGTTCTTACATCAGATAGTGGTTTCAGATTTACGCGAGTGAATTCAGATTTCAGC 2099
Db 2041 TTGCGGTAGTAGTTCTTACATCAGATAGTGGTTTCAGATTTACGCGAGTGAATTCAGATTTCAGC 2100
Qy 2100 AAGTGATTTGAGCTTCAGCGAGTGATTCAGATTTACGAAAGCGATTTCCGACTCAGCGAGCGA 2159
Db 2101 AAGTGATTTGAGCTTCAGCGAGTGATTCAGATTTACGAAAGCGATTTCCGACTCAGCGAGCGA 2160
Qy 2160 TTCCGACTCAGACAAATGACTCGGATTCAGATTTACGATAGCGATTTCTGACTCAGACAGTACTCAGA 2219
Db 2161 TTCCGACTCAGACAAATGACTCGGATTCAGATTTACGATAGCGATTTCTGACTCAGACAGTACTCAGA 2220
Qy 2220 TTCCGACAGTGCAGATTCAGATTTACGATAGCGATTTCTGACTCAGACAGTACTCAGA 2279
Db 2221 TTCCGACAGTGCAGATTCAGATTTACGATAGCGATTTCTGACTCAGACAGTACTCAGA 2280
Qy 2280 TAGCGATTTGAGTTTCAGATTTAGCGATTCAGATTTCCGACAGTGAATTCGACTCAGACAGCGA 2339
Db 2281 TAGCGATTTGAGTTTCAGATTTAGCGATTCAGATTTCCGACAGTGAATTCGACTCAGACAGCGA 2340
Qy 2340 TTCTGACTCCGACAGTGAATTCGACTCAGACAGGATTCAGATTTCCGACAGTGAATTCGGA 2399
Db 2341 TTCTGACTCCGACAGTGAATTCGACTCAGACAGGATTCAGATTTCCGACAGTGAATTCGGA 2400
Qy 2400 CTGAGATAGCGATTTCCGACTCAGATAGCGACTCAGATTTCCGACAGTGAATTCAGATTTCCGA 2459
Db 2401 CTGAGATAGCGATTTCCGACTCAGATAGCGACTCAGATTTCCGACAGTGAATTCAGATTTCCGA 2460
Qy 2460 CAGCGATTTGAGTTTCAGATTTAGCGATTCAGATTTCCGACAGTGAATTCGACTCAGACAGTGA 2519
Db 2461 CAGCGATTTGAGTTTCAGATTTAGCGATTCAGATTTCCGACAGTGAATTCGACTCAGACAGTGA 2520
Qy 2520 CTGCGATTTGAGTGAATTCGAGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2579
Db 2521 CTGCGATTTGAGTGAATTCGAGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2580
Qy 2580 CTGAGACAGTGAATTCGAGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2639
Db 2581 CTGAGACAGTGAATTCGAGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2640
Qy 2640 CAGTGACTCGGATTCAGATTTAGCGACTCAGATTTCCGACAGTGAATTCGATTTCCGACAGTGA 2699
Db 2641 CAGTGACTCGGATTCAGATTTAGCGACTCAGATTTCCGACAGTGAATTCGATTTCCGACAGTGA 2700
Qy 2700 TTCCGACTCAGATAGCGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2759
Db 2701 TTCCGACTCAGATAGCGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2760
Qy 2760 TTCCGACTCAGATAGCGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2819
Db 2761 TTCCGACTCAGATAGCGATTTCCGACAGTGAATTCGATTTCCGACAGTGAATTCGGA 2820

Db	6447	TGAAAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTC	6388	Qy	1560	GGATAGTAATGCATTAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAA	1619
Qy	480	TAGCGTTAGTCTGCACCTAAACACAGACGACACAAAACGTGAGTGATACCTAAAAACATCGTC	539	Db	5307	GGATAGTAATGCATTAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAA	5248
Db	6387	TAGCGTTAGTCTGCACCTAAACACAGACGACACAAAACGTGAGTGATACCTAAAAACATCGTC	6328	Qy	1620	TGCAGCTGATTTATCTGAAAGTTACTTTGTGGAATCCAGAAACCTTTGAGGATGTCACATA	1679
Qy	540	AAACACTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGACACAAAGGAAACGACACA	599	Db	5247	TGCAGCTGATTTATCTGAAAGTTACTTTGTGGAATCCAGAAACCTTTGAGGATGTCACATA	5188
Db	6327	AAACACTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGACACAAAGGAAACGACACA	6268	Qy	1680	TAGTGTGAATTAATACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTATACGCTGA	1739
Qy	600	ATCATCATCAAAATGCACTAGCGAAGAAACCGCGGTAACTGGTGAAGCTACTACTAC	659	Db	5187	TAGTGTGAATTAATACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTATACGCTGA	5128
Db	6267	ATCATCATCAAAATGCACTAGCGAAGAAACCGCGGTAACTGGTGAAGCTACTACTAC	6208	Qy	1740	TGATCAAAATTAACAACCGGTATATAGTAGTTGTTTAATGGTCAATATGATCCGAATAGCAA	1799
Qy	660	GACACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGGGAGGA	719	Db	5127	TGATCAAAATTAACAACCGGTATATAGTAGTTGTTTAATGGTCAATATGATCCGAATAGCAA	5068
Db	6207	GACACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGGGAGGA	6148	Qy	1800	AGGTGATTTAGCTTTTACGTTCAACTTTATATGGGTATAACTCGAATATAATTTGGCGCTC	1859
Qy	720	ATTAGTGAATCAAAAGTAATGAACGACTTTTAAATGATACTAATACAGTATCATCTGT	779	Db	5067	AGGTGATTTAGCTTTTACGTTCAACTTTATATGGGTATAACTCGAATATAATTTGGCGCTC	5008
Db	6147	ATTAGTGAATCAAAAGTAATGAACGACTTTCTAATGATACTAATACAGTATCATCTGT	6088	Qy	1860	TATGTCATGGGACAAACGAGTAGCATTTAATAACGGATCAGGTTCTGCTGACGCTATCGA	1919
Qy	780	AAATTCACCTCAAAATTTACAAATGGGAAATGTTTCAACACGCGAGATCTTCAAC	839	Db	4947	TAAACCAAGTTGTTCTTCAACCAACCTGATGAGCTTGGTGAATTTGAACCAATTCAGAGGA	4888
Db	6087	AAATTCACCTCAAAATTTACAAATGGGAAATGTTTCAACACGCGAGATCTTCAAC	6028	Qy	1920	TAAACCAAGTTGTTCTTCAACCAACCTGATGAGCTTGGTGAATTTGAACCAATTCAGAGGA	1979
Qy	840	TGAAGCAACACCTTCAAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGA	899	Db	4947	TAAACCAAGTTGTTCTTCAACCAACCTGATGAGCTTGGTGAATTTGAACCAATTCAGAGGA	4888
Db	6027	TGAAGCAACACCTTCAAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGA	5968	Qy	1980	TTGAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGGATTTCAATTCAGATAGCGGTTTCA	2039
Qy	900	TGTAGTTAATCAACGGGTTAATAAAGTGCGCTTAGAATGAGACATTTAGTTAGCGGC	959	Db	4887	TTGAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGGATTTCAATTCAGATAGCGGTTTCA	4828
Db	5967	TGTAGTTAATCAACGGGTTAATAAAGTGCGCTTAGAATGAGACATTTAGTTAGCGGC	5908	Qy	2040	TTCCGGTAGTGATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTATTCAGATTTCAGC	2099
Qy	960	AGTAGCTGCAGATGACCGGACGCTGCGACAGATATTACGATCAGTTGACGATGTGAC	1019	Db	4827	TTCCGGTAGTGATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTATTCAGATTTCAGC	4768
Db	5907	AGTAGCTGCAGATGACCGGACGCTGCGACAGATATTACGATCAGTTGACGATGTGAC	5848	Qy	2100	AAAGTATTTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	2159
Qy	1020	AGTTGGTATTGACCTCTGGTACGATCTGATCCGACCAACGAGGTTATGTCAAACTGAA	1079	Db	4767	AAAGTATTTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	4708
Db	5847	AGTTGGTATTGACCTCTGGTACGATCTGATCCGACCAACGAGGTTATGTCAAACTGAA	5788	Qy	2160	TTCCGACTCAGCAATGACTCGGATTCAGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	2219
Qy	1080	TTATGGTTTTTCACTGCTGCTAATCTGCTGTTTAAAGGTGACACATTCAAAATGATACC	1139	Db	4707	TTCCGACTCAGCAATGACTCGGATTCAGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	4648
Db	5787	TTATGGTTTTTCACTGCTGCTAATCTGCTGTTTAAAGGTGACACATTCAAAATGATACC	5728	Qy	2220	TTCCGACTCAGCAATGACTCGGATTCAGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	2279
Qy	1140	TAAAGATTAACCTTAATGGTGTAACTTCACTGCTTAAAGTGCACCAATATGGCTGG	1199	Db	4647	TTCCGACTCAGCAATGACTCGGATTCAGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	4588
Db	5727	TAAAGATTAACCTTAATGGTGTAACTTCACTGCTTAAAGTGCACCAATATGGCTGG	5668	Qy	2280	TAGCGATTTCAGATTTCAGATTCAGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	2339
Qy	1200	AGATCAAGTATTGGCAATGGTGTAAATCGATAGTGGTAAATGTTATTTATACATTTAC	1259	Db	4587	TAGCGATTTCAGATTTCAGATTCAGATTCAGATTCAGCAAGCGGATTCGCACTCAGCGAGCGA	4528
Db	5667	AGATCAAGTATTGGCAATGGTGTAAATCGATAGTGGTAAATGTTATTTATACATTTAC	5608	Qy	2340	TTCTGACTCCGACAGTGAATTCGCACTCAGACAGCGGATTCAGATTCGCAAGCGGATTCGCA	2399
Qy	1260	AGACTATGTAATACTAAGATGATGTAAGAGAACCTTTGACCAAGCGCGCTTATATGA	1319	Db	4527	TTCTGACTCCGACAGTGAATTCGCACTCAGACAGCGGATTCAGATTCGCAAGCGGATTCGCA	4468
Db	5607	AGACTATGTAATACTAAGATGATGTAAGAGAACCTTTGACCAAGCGCGCTTATATGA	5548	Qy	2400	CTCAGATAGCGATTTCGCACTCAGATTCAGATTCAGATTCAGCAAGCGGATTCAGATTCAGCA	2459
Qy	1320	CCCTGAAAACTTAAAGACAGGTAAATGTCACATTTGCTACTGCTAGGTAGTACCAAC	1379	Db	4467	CTCAGATAGCGATTTCGCACTCAGATTCAGATTCAGATTCAGCAAGCGGATTCAGATTCAGCA	4408
Db	5547	CCCTGAAAACTTAAAGACAGGTAAATGTCACATTTGCTACTGCTAGGTAGTACCAAC	5488	Qy	2460	CAGCGATTTCAGATTTCAGATTCAGATTCAGATTCAGCAAGCGGATTCAGATTCGCAAGCGGATTCAGATTCAGCA	2519
Qy	1380	AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAAGTTTATATCTTCTAT	1439	Db	4407	CAGCGATTTCAGATTTCAGATTCAGATTCAGATTCAGCAAGCGGATTCAGATTCGCAAGCGGATTCAGATTCGCA	4348
Db	5487	AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAAGTTTATATCTTCTAT	5428	Qy	2520	CTCGGATTTCAGATTCAGATTCAGATTCAGATTCAGCAAGCGGATTCAGATTCGCAAGCGGATTCAGATTCGCA	2579
Qy	1440	TAAAGGTACAAATTCGCAAAATCGATATAAACHAATTAATACGATTCGTCAGACAAATTTATGT	1499	Db	4347	CTCGGATTTCAGATTCAGATTCAGATTCAGATTCAGCAAGCGGATTCAGATTCGCAAGCGGATTCAGATTCGCA	4288
Db	5427	TAAAGGTACAAATTCGCAAAATCGATATAAACHAATTAATACGATTCGTCAGACAAATTTATGT	5368	Qy	2580	CTCAGACAGTGAATTCGCGATTTCAGCGAGTGAATTCGCGATTTCAGATAGTGAATTCGCGACTCCGA	2639
Qy	1500	CAATCCAAAGTGGAGATTAACGTTATTTGGCGCGGTTTTTAAACAGGTAATTTAAACCAAAATAC	1559	Db	4287	CTCAGACAGTGAATTCGCGATTTCAGCGAGTGAATTCGCGATTTCAGATAGTGAATTCGCGACTCCGA	4228
Db	5367	CAATCCAAAGTGGAGATTAACGTTATTTGGCGCGGTTTTTAAACAGGTAATTTAAACCAAAATAC	5308				

QY 2640 CAGTACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 2699
DB 4227 CAGTACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 4168
QY 2700 TTCGGACTCAGATAGCGATTCAGAAATCAGACGCGATTCAGAAATCAGACGCGATTCAGA 2759
DB 4167 TTCGGACTCAGAT-----AGCGATTCAGAAATCAGACGCGATTCAGA 4126
QY 2760 TTCAGACGCGACTCAGACGAGTGCATCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2819
DB 4125 TTCAGACGCGACTCAGACGAGTGCATCAGATTCAGATTCAGATTCAGATTCAGATTCAG 4066
QY 2820 TTCAGACTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2879
DB 4065 TTCAGACTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 4006
QY 2880 TAGCGATTCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2939
DB 4005 TAGCGATTCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 3946
QY 2940 TACTAATGCTTCTAATAAATAGAGCTAAGATAGTAAAGACCAATACCGATACAGG 2999
DB 3945 TACTAATGCTTCTAATAAATAGAGCTAAGATAGTAAAGACCAATACCGATACAGG 3886
QY 3000 TTCTGAAGATGAAGCAATACGTCACCTAAATTTGGGATTCATAGCATCAATAGGTTCAAT 3059
DB 3885 TTCTGAAGATGAAGCAATACGTCACCTAAATTTGGGATTCATAGCATCAATAGGTTCAAT 3826
QY 3060 ACTACTTTTTCAG 3119
DB 3825 ACTACTTTTTCAG 3766
QY 3120 TTAATCATATGATTCATGAAG 3179
DB 3765 TTAATCATATGATTCATGAAG 3706
QY 3180 AAATATATGTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3239
DB 3705 AAATATATGTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3646
QY 3240 GTAGATGTTTATATAATTTGGCTTTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3299
DB 3645 GTAGATGTTTATATAATTTGGCTTTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3586
QY 3300 GGAAATTAAG 3359
DB 3585 GGAAATTAAG 3526
QY 3360 TATGTGTATCAGAAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3419
DB 3525 TATGTGTATCAGAAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3466
QY 3420 ATCCATATGATCTGAGTTCGTTGAAATGACTAATAATAATAATAATAATAATAATAATA 3479
DB 3465 ATCCATATGATCT-AGTGTGAAATGACTAATAATAATAATAATAATAATAATAATAATA 3407
QY 3480 AAGAGTATAGAAAAGCTT 3498
DB 3406 AAGAGTATAGAAAAGCTT 3388

RESULT 5
AR354469
LOCUS
DEFINITION
Accession
AR354469
VERSION
AR354469.1
KEYWORDS
GI:33760553
SOURCE
Unknown.
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 4709)
AUTHORS
Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and

Rosen,C.A.
Staphylococcus aureus polynucleotides and sequences
Patent: US 6593114-A 587 15-JUL-2003;
Location/Qualifiers
1..4709
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 97.2%; Score 3399.6; DB 6; length 4709;
Best Local Similarity 99.1%; Pred.No.0;
Matches 3466; Conservative 1; Mismatches 11; Indels 21; Gaps 4;

QY 1 GGTACCAATAAATTCACATCTCTCTTTTGAAAAATATGATTTCAAGCTAGGATTCATTA 60
DB 1204 GGTACCAATAAATTCACATCTCTCTTTTGAAAAATATGATTTCAAGCTAGGATTCATTA 1263
QY 61 GGTAGAGTTTCATATTAATAATAAATAAATGTTTGCAATCAATCGTACGTTGTCGTTGTA 120
DB 1264 GGTAGAGTTTCATATTAATAATAAATAAATGTTTGCAATCAATCGTACGTTGTCGTTGTA 1323
QY 121 ATTCTTAATAAGCAATAAATAAATAAATGTTTGTAAGATATTTGTAAGATATTTGTAAGAT 180
DB 1324 ATTCTTAATAAGCAATAAATAAATAAATGTTTGTAAGATATTTGTAAGATATTTGTAAGAT 1383
QY 181 TATCGATACAAATTAATTCCTATATATGCAATTTTAGTGATATTAATTCATTAACAGAGATT 240
DB 1384 TATCGATACAAATTAATTCCTATATATGCAATTTTAGTGATATTAATTCATTAACAGAGATT 1443
QY 241 AAATATATC-TTAAAGGGTATATAGTTTAATAATAAATGACTTTTAAAAAGAGGGGAATAA 299
DB 1444 AAATATATC-TTAAAGGGTATATAGTTTAATAATAAATGACTTTTAAAAAGAGGGGAATAA 1503
QY 300 AATGAATATGAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 359
DB 1504 AATGAATATGAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1563
QY 360 AGTGTCTGTAGGTAGCTTAATCGGTTTTCGACTACTCAGCAGTAAAGAGAGAGATCAAG 419
DB 1564 AGTGTCTGTAGGTAGCTTAATCGGTTTTCGACTACTCAGCAGTAAAGAGAGAGATCAAG 1623
QY 420 TGAATAATGTTTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAGCAAGCAAGCAAGCAAG 479
DB 1624 TGAATAATGTTTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAGCAAGCAAGCAAGCAAG 1683
QY 480 TAGCGTTAGTGTGCACTTAAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 539
DB 1684 TAGCGTTAGTGTGCACTTAAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 1743
QY 540 AAACACTTAATAATGGCGAAACGAGTGTGGCGAAATCCAGCAACAGCAACAGCAACAGCAACAG 599
DB 1744 AAACACTTAATAATGGCGAAACGAGTGTGGCGAAATCCAGCAACAGCAACAGCAACAGCAACAG 1803
QY 600 ATCATCATCAACAAATGCAACTACCGAAGAAACGCGGTAACTGGTGAAGTACTACTACTAC 659
DB 1804 ATCATCATCAACAAATGCAACTACCGAAGAAACGCGGTAACTGGTGAAGTACTACTACTAC 1863
QY 660 GACACGAATCAAGCTAATAATACCGCGCAACAACTCAATCAAGCAATCAAGTCCGAGGA 719
DB 1864 GACACGAATCAAGCTAATAATACCGCGCAACAACTCAATCAAGCAATCAAGTCCGAGGA 1923
QY 720 ATTAGTGAATCAACAAAGTAAATGAACGACTTTTATGATATTAATCAAGTATCACTCTGT 779
DB 1924 ATTAGTGAATCAACAAAGTAAATGAACGACTTTTATGATATTAATCAAGTATCACTCTGT 1983
QY 780 AAATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACAAACGCAAGATCTTCAAC 839
DB 1984 AAATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACAAACGCAAGATCTTCAAC 2043
QY 840 TGAAGCAACACCTTCAACAAATGAATGAGTCCAGAGATACAGATGCAAGTAAATTAAGA 899
DB 2044 TGAAGCAACACCTTCAACAAATGAATGAGTCCAGAGATACAGATGCAAGTAAATTAAGA 2103

QY 2640 CAGTACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 2699
DB 4227 CAGTACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 4168
QY 2700 TTCGGACTCAGATAGCGATTCAGAAATCAGACGCGATTCAGAAATCAGACGCGATTCAGA 2759
DB 4167 TTCGGACTCAGAT-----AGCGATTCAGAAATCAGACGCGATTCAGA 4126
QY 2760 TTCAGACGCGACTCAGACGAGTGCATCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2819
DB 4125 TTCAGACGCGACTCAGACGAGTGCATCAGATTCAGATTCAGATTCAGATTCAGATTCAG 4066
QY 2820 TTCAGACTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2879
DB 4065 TTCAGACTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 4006
QY 2880 TAGCGATTCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2939
DB 4005 TAGCGATTCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 3946
QY 2940 TACTAATGCTTCTAATAAATAGAGCTAAGATAGTAAAGACCAATACCGATACAGG 2999
DB 3945 TACTAATGCTTCTAATAAATAGAGCTAAGATAGTAAAGACCAATACCGATACAGG 3886
QY 3000 TTCTGAAGATGAAGCAATACGTCACCTAAATTTGGGATTCATAGCATCAATAGGTTCAAT 3059
DB 3885 TTCTGAAGATGAAGCAATACGTCACCTAAATTTGGGATTCATAGCATCAATAGGTTCAAT 3826
QY 3060 ACTACTTTTTCAG 3119
DB 3825 ACTACTTTTTCAG 3766
QY 3120 TTAATCATATGATTCATGAAG 3179
DB 3765 TTAATCATATGATTCATGAAG 3706
QY 3180 AAATATATGTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3239
DB 3705 AAATATATGTTTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3646
QY 3240 GTAGATGTTTATATAATTTGGCTTTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3299
DB 3645 GTAGATGTTTATATAATTTGGCTTTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3586
QY 3300 GGAAATTAAG 3359
DB 3585 GGAAATTAAG 3526
QY 3360 TATGTGTATCAGAAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3419
DB 3525 TATGTGTATCAGAAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3466
QY 3420 ATCCATATGATCTGAGTTCGTTGAAATGACTAATAATAATAATAATAATAATAATAATA 3479
DB 3465 ATCCATATGATCT-AGTGTGAAATGACTAATAATAATAATAATAATAATAATAATAATA 3407
QY 3480 AAGAGTATAGAAAAGCTT 3498
DB 3406 AAGAGTATAGAAAAGCTT 3388

RESULT 5
AR354469
LOCUS
DEFINITION
Accession
AR354469
VERSION
AR354469.1
KEYWORDS
GI:33760553
SOURCE
Unknown.
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 4709)
AUTHORS
Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and

after

Db 4246 ACTACTTTTCAGAGAAAAAGAAAAATAAGATAGAAAAATAAGTAATAATCATATTAAA 4305
QY 3120 TTAATCATATGATTCATGAGAGAGCCACCTTAAAGGTCCTTCTTTTACTTGGATTTCC 3179
Db 4306 TTAATCATATGATTCATGAGAGAG-ACRCCTTAAAGGTCCTTCTTTTACTTGGATTTCC 4364
QY 3180 AAATATATTTGTTGAATATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 3239
Db 4365 AAATATATTTGTTGAATATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 4424
QY 3240 GTAGATGTTATATAATTTGGCTTGGCGAAGAAATAGGGTGTAGGTAGGTGTTAATTAAG 3299
Db 4425 GTAGATGTTATATAATTTGGCTTGGCGAAGAAATAGGGTGTAGGTAGGTGTTAATTAAG 4484
QY 3300 GGAATTAAGAGAGAAATACAGTTCAAAATAATAATTTGCTAGTTTATCATTTGGGAGCAT 3359
Db 4485 GGAATTAAGAGAGAAATACAGTTCAAAATAATAATTTGCTAGTTTATCATTTGGGAGCAT 4544
QY 3360 TATGTGTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTCTGGGAGAGAGA 3419
Db 4545 TATGTGTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTCTGGGAGAGAGA 4604
QY 3420 ATCCATATGATCTGAGTGGTGAAGTCACTGAAATTAATAATAATAATAATAATAATAATA 3479
Db 4605 ATCCATATGATCT-AGTGGTTGAAACTGACTAATAATAATAATAATAATAATAATAATA 4663
QY 3480 AAGAGTATAAGAAAGCTT 3498
Db 4664 AAGAGTATAAGAAAGATT 4682

RESULT 6

BX571857_08

WPCOMMENT

Sequence split into 28 fragments LOCUS BX571857 Accession BX571857

Fragment Name	Begin	End
BX571857_00	1	110000
BX571857_01	100001	210000
BX571857_02	200001	310000
BX571857_03	300001	410000
BX571857_04	400001	510000
BX571857_05	500001	610000
BX571857_06	600001	710000
BX571857_07	700001	810000
BX571857_08	800001	910000
BX571857_09	900001	1010000
BX571857_10	10000001	11100000
BX571857_11	11000001	12100000
BX571857_12	12000001	13100000
BX571857_13	13000001	14100000
BX571857_14	14000001	15100000
BX571857_15	15000001	16100000
BX571857_16	16000001	17100000
BX571857_17	17000001	18100000
BX571857_18	18000001	19100000
BX571857_19	19000001	20100000
BX571857_20	20000001	21100000
BX571857_21	21000001	22100000
BX571857_22	22000001	23100000
BX571857_23	23000001	24100000
BX571857_24	24000001	25100000
BX571857_25	25000001	26100000
BX571857_26	26000001	27100000
BX571857_27	27000001	27998002

Continuation (9 of 28) of BX571857 from base 800001 (BX571857 Staphylococcus aureus stra

Query Match 93.7%; Score 3277; DB 1; Length 110000;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 3383; Conservative 0; Mismatches 100; Indels 16; Gaps 3;

QY 1 GGTACCATAAATACACATCTGCTTTTGAAAAAATATGATTTCAGAGTAGGATTACATTA 60

Db 40748 GGTACCATAAATACACATCTGCTTTTGAAAAAATATGATTTCAGAGTAGGATTACATTA 40807

QY 61 GGTAGAGTTTCATTTAATAATAAAAAATGTTTGCATCAATCAATCGTACGTTGCTGTTGTA 120
Db 40808 GGTAGAGTTTCATTTAATAATAAAAAATGTTTGCATCAATCGTACGTTGCTGTTGTA 40867
QY 121 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 180
Db 40868 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 40927
QY 181 TATCGATACAAATTAATTTGCTATATGCAATTTTAGTGTATATTAATTCATTAACAGAGATT 240
Db 40928 TATCGATACAAATTAATTTGCTATATGCAATTTTAGTGTATATTAATTCATTAACAGAGATT 40987
QY 241 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTAAAGAGGGGAATAA 299
Db 40988 AAATATATCTTTAAAGGGTATATAGTTAATAATAAATGACTTTTAAAGAGGGGAATAA 41047
QY 300 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC 359
Db 41048 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC 41107
QY 360 AGTGCCTTTAGGTACGTTAATCGTTTTTGGACTTACTCAGCAGTAAAGAACGATGCAAG 419
Db 41108 AGTGCCTTTAGGTACGTTAATCGTTTTTGGACTTACTCAGCAGTAAAGAACGATGCAAG 41167
QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTCAAG 479
Db 41168 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGTAATGATTCAAG 41227
QY 480 TAGCGTTAGTGTGCACCTTAAACACAGACACAAACGTTGAGTGTATCTAATAACATCGTC 539
Db 41228 TAGCGTTAGTGTGCACCTTAAACACAGACACAAACGTTGAGTGTATCTAATAACATCGTC 41287
QY 540 AAACACTAATATGGCGAAACGAGTGTGGCGGAAATCCAGACACAAACGAGTAAACACGACA 599
Db 41288 AAACACTAATATGGCGAAACGAGTGTGGCGGAAATCCAGACACAAACGAGTAAACACGACA 41347
QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAAGCGCGGTAACTGTGTGAAGTACTACTAC 659
Db 41348 ATCAGCNTTAAACAAATGCAACTACGGAAGAAAGCGCGGTAACTGTGTGAAG--CTACTAC 41404
QY 660 GACAAAGTAAGTGTATATACACCGGCAACAACTCAATCAAGCAATCAAAATCGCGAGGA 719
Db 41405 GCAACGTAAGTGTATATACACCGGCAACAACTCAATCAAGCAATCAAAATCGCGAGGA 41464
QY 720 ATTAGTGAATCAACAAAGTAATGAACGACTTTTATGATCTACTAATACAGTATCATCTGT 779
Db 41465 ATTAGTGAATCAACAAAGTAATGAACGACTTTTATGATCTACTAATACAGTATCATCTGT 41524
QY 780 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATACCTTCAAC 839
Db 41525 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATACCTTCAAC 41584
QY 840 TGAAGCAACACCTTCAACAAATGAATCAGTCCACAGAGTACAGATGCAAGTAAATAAGA 899
Db 41585 TGAAGCAACACCTTCAACAAATGAATCAGTCCACAGAGTACAGATGCAAGTAAATAAGA 41644
QY 900 TGTAGTTAATCAAGCGGTTTAAATACAGTGGCGCTAGAATGAGAGCATTTTAGTTTACGGC 959
Db 41645 TGTAGTTAATCAAGCGGTTTAAATACAGTGGCGCTAGAATGAGAGCATTTTAGTTTACGGC 41704
QY 960 AGTAGTCGAGATGCAACCGGAGTGGCAGAGATATTACGAATCAGTTGACGAATGTGAC 1019
Db 41705 AGTAGTCGAGATGCAACCGGAGTGGCAGAGATATTACGAATCAGTTGACGAATGTGAC 41764
QY 1020 AGTTGTTATTGACTCTGGTACCACTGTGTATCGCACCAAGCAGGTTATGTCAAACGTAA 1079
Db 41765 AGTTGTTATTGACTCTGGGAGATACAGTTTATCGCACCAAGCAGGCTATGTCAAACGTAA 41824
QY 1080 TTATGTTTTTTCAGTGCCTTAATCTGCTGTTAAAGTGACACATTCAAAATAACTGTACC 1139
Db 41825 TTATGTTTTTTCAGTGCCTTAATCTGCTGTTAAAGTGACACATTCAAAATAACTGTACC 41884

Db 44033 CGAARATTAGGAGAAATACAGTTGAAATAATAATTGCTAGTCTTTATCATTTGGGAGCAT 44092
QY 3360 TAATGTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTCTTCTGGGAGAGA 3419
Db 44093 TATGTATCAAAATTTGGGAAAGTAATCGTGGAGTGCAGTGGTTCTTCTGGGAGAGA 44152
QY 3420 ATCCATATGATCTAGTCTGCTGAACTGACTAATAATAATAATAATACTAGAACAGTAG 3479
Db 44153 ATCCATATGATCTAGTCTGCTGAACTGACTAATAATAATAATACTAGAACAGTAG 44212
QY 3480 AAGAGTATAAGAAAGCTT 3498
Db 44213 AAGAGTATAAGAAAGCTT 44231

RESULT 7
AP004824 290150 bp DNA linear BCT 20-DEC-2002
LOCUS
DEFINITION
Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain:MW2, section 3/10.
ACCESSION
AP004824 BA000033
VERSION
AP004824.1 GI:21203693
KEYWORDS
SOURCE
Staphylococcus aureus subsp. aureus MW2
ORGANISM
Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1
Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A.,
Nagai, Y., Iwama, N., Asano, K., Nakamura, T., Kuroda, H., Cui, L.,
Yamamoto, K. and Hiratsuka, K.
Genome and virulence determinants of high virulence
community-acquired MRSA
Lancet 359 (9320), 1819-1821 (2002)
1204378
2 (bases 1 to 290150)
Director-General Biotechnology Center, Aoki, K., Oguchi, A.,
Nagai, Y., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiratsuka, K. and
Kikuchi, H.
Direct Submission
Submitted (06-MAR-2002) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center, A-cho 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
E-mail: bio.nite.go.jp, URL: http://www.bio.nite.go.jp/,
Tel: 81-3-3481-1933, Fax: 81-3-3481-8424
FEATURES
Location/Qualifiers
1. 290150
/organism="Staphylococcus aureus subsp. aureus MW2"
/mol_type="genomic DNA"
/strain="MW2"
/db_xref="taxon:196620"
345..1745
/gene="prop"
345..1745
/gene="prop"
/note="ORFID:MW0528
proline/betaine transporter homologue"
/codon_start=1
/transl_table=11
/protein_id="BAB94393.1"
/db_xref="GI:21203694"
/translation="MDENKENINWDAKAKKATVATGIGNAMEWDFGVYATYAVI
GANFFSVENADIQMLTFAALAFLLRPIGGVGVVGLIGDKYGRKVVLSITLLMAP
STITGLPSYDQIGLWAPLILLARLPQSTGCGAGATVVAESPKRRNSLGS
GLEITGLSYGTAIAVIAVLTFFLTDDQASFGWRIPFLGLFLGLFLYLRRLERS
PVENDVATPERDINFLQIRFYIDIFVCFVAVFVFNVTNMTAYLPTYLEQVI
KLDATTSVLITCWAIMIPALMFKGLADKIGKKVFLTGGLTLFSLIAFWLLHS
QSFVIVIGFILGFUSTYEATMPGSLPTMFYSHIRYRILSVTFNLSVSIFGTTPL
VATLVTKGDPLAPAYLLTAISVIFLVTIFLHSTAGSKLSGSPNVNDNQDRAV
AEHPKEALMWYKERKN"
2080..2289
/gene="MW0529"
2080..2289

gene
CDS
2080..2289
/gene="MW0529"
2080..2289

gene
CDS
2289..3665
/gene="vraA"
2289..3665
/gene="vraA"
/note="ORFID:MW0530
hypothetical protein, similar to long chain fatty acid CoA
ligase"
/codon_start=1
/transl_table=11
/protein_id="BAB94395.1"
/db_xref="GI:21203696"
/translation="MNVILEQLTKHTQNPNDIALHIDDEITYTYSOLNARITSAVESL
QKSLNPVAINMKSPVQSIICYLALHRLHKVMMMEGKQSTIHRQLIEKYIKEDVI
GUTGLMONIDSEMFIDSTOLQHPNLLHIGFTSGTGLPKAYRDEDSWLASFVNEM
LMKNENIAAPGUSHSLTLXALLFALSSGRIFGQTFHFEKLLNQCHKISSYKVA
MFLVPTMKSLLLVYNNHEITQSFSSGDKLHSSIPKIKNQKANDLNLIEFGTSETIS
FISYLNQAPVESYGLFPNVELKLTNHDHNGIGTICIKSNMFGSYVSEQCINND
WFVNDNGYVQSYLYLTGRQODMLIGQNTYPAHVERLLTQSSSIDAIIIGIPNE
RFGQGLVLLYSGDVLTHTKNVQFLKKVKRYEIPSMIHVEKMYVTASGKIAREKMM
SMILRGEEL"
3667..4806
/gene="vraB"
3667..4806
/gene="vraB"
/note="ORFID:MW0531"
/codon_start=1
/transl_table=11
/product="acetyl-CoA c-acetyltransferase"
/protein_id="BAB94396.1"
/db_xref="GI:21203697"
/translation="MNOAVIVAARAKTAFGKYGTGLKHLRPEQLLKPLFOHFKEKYPEV
ISKDIDVILGNVVGNGNARKALLEAGLKDSIPGVITDRQCSGLSEYQYACRMICA
GKQVIAGVVESTSPAPWKIKRPHSYETALPEFYERASFAPEDSPMIQAEVNA
KDYDREQLDEFAYRSHQLTAEVKNNGNISQELPITVKGEIFNDESLKSHIPDN
KRFKFKVIGGTVAANSCKMNDGAVLLIMEKDMAYELGFEHGLFKDGVTVGVDSN
FPGIGPVPAISLLKENOLTENIEVIEINEAFSAQVACQALNISTQNLNIGGAL
ASCHPYGASGAGLVTLFLYMFEDKETWIASNGIGGGLGNAALFTRF"
4781..5146
/gene="vraC"
4781..5146
/gene="vraC"
/note="ORFID:MW0532"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB94397.1"
/db_xref="GI:21203698"
/translation="MQHLLDSNQRLNVSPSKDSVAAYOCFNQPYRKEYVPLMCA
WPKEDLFKXANSELILTKSAINQTKQIEVDITVUGHLEDIECRQENIRYTMALTL
TKNDQHVIVITQIFIKAMK"
5149..5418
/gene="MW0533"
5149..5418
/gene="MW0533"
/note="ORFID:MW0533"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB94398.1"
/db_xref="GI:21203699"
/translation="MKRFENIWEINNEYALVNDNDPIHNEIVPGQLVSOYMLMAMSL
EQCQIVYKPLINENIEFEQHEHIIAINDDGEIKIKISLSTKK"
complement (5619..5786)
/gene="MW0534"

gene
CDS

CDS	complement(5619..5786) /gene="MW0534" /note="ORFID:MW0534" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94399.1" /db_xref="GI:21203700" /translation="MIIYRQVHHEGAPVVEITKTFQHSIKCDDSFSDTEIFKLKLSLQDDIDHMKVS" complement(6293..7123) /gene="MW0535" complement(6293..7123) /gene="MW0535" /note="ORFID:MW0535" hypothetical protein, similar to phosphomethylpyrimidine kinase, thid homologue" /codon_start=1 /transl_table=11 /protein_id="BAB94400.1" /db_xref="GI:21203701" /translation="MALKKVLTIAGSDTSAGAGMADLKTFOELDTYGMVALTAVTMKDQWNSHDVTPLEMDVFEKOLETALSIGPDALKTGMGTSEIIRAGEVYEASNAQYFVVDPMVCKGEDEVINGENTAMIKYLLPKAIYVTPNLFAGQSLGKUNSIEDMKKAAATLIFDKGAQHVLIKGGKALDQKSDLYDYDQGYQLTTFDMFQQSYNHGAGCTFAAATTAYLANGSPKPEAVISAKAFVASAIKNGWKMDVFGVDHGYNRIEIHIDVEVTEV" 7307..7963 /gene="ung" 7307..7963 /gene="ung" /note="ORFID:MW0536" /codon_start=1 /transl_table=11 /product="uracil-DNA glycosylase" /protein_id="BAB94401.1" /db_xref="GI:21203702" /translation="MENSQIFPHDITTKHDFKAMHDFLEKEYSTAIVYDPDRENIYQAFDLTPENIKVILGDDPQHQAGLAFSPQVNPNAKFPSPSLRNMYKEHLADIGCVROTPHLQDWARREGLLNTVLTVROGEANSRHDIGWETFDRIKAVSDYKEHVYVILWGKFAQQKIKLIDTSKHCIKIISVHPSPLSAYRGFGSKPYSKANNAYLSVGKSPINWCESEA" 7964..8344 /gene="MW0537" 7964..8344 /gene="MW0537" /note="ORFID:MW0537" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="BAB94402.1" /db_xref="GI:21203703" /translation="MLARETLIARIEQLVQAEQAOHQHDDPEKHMVYAIHLITSLYASTSNTPHLGEQWNRRIANHNQMPQSIQTQPHQVTALEIANGGKVNTHSAHHHNKSYSQPSQQORLATDDDIGNGESIFDF" 8477..8845 /gene="MW0538" 8477..8845 /gene="MW0538" /note="ORFID:MW0538" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="BAB94403.1" /db_xref="GI:21203704" /translation="MKLFIILGALNNMAYGTGAFGAHGLQGLKISDRHLSYWEKATYYQMVFLLGIWMLIATPFKFAF" 8966..10450 /gene="MW0539" 8966..10450	Best Local Similarity 95.8%; Pred. No. 0; Matches 3403; Conservative 0; Mismatches 80; Indels 70; Gaps 4;
gene	1 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTACATTA 60 Db 241201 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGCTAGGATTACATTA 241260	
gene	61 GGTAGAGTTCATATTAATAAATAAATGTTTGCATCAAAATCGACGTTGCGTTGTA 120 Db 241261 GGTAGAGTTCATATTAATAAATAAATGTTTGCATCAAAATCGACGTTGCGTTGTA 241320	
gene	121 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTCTGGATAATAAAA 180 Db 241321 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTCTGGATAATAAAA 241380	
gene	181 TATCGATACAAATTAATTTGCTATATGCAATTTTAGTGTATAATTCCTATTAACAGAGATT 240 Db 241381 TATCGATACAAATTAATTTGCTATATGCAATTTTAGTGTATAATTCCTATTAACAGAGATT 241440	
gene	241 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTTTAAAGAGGGAATAA 299 Db 241441 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTTTAAAGAGGGAATAA 241500	
gene	300 AATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGCCTTC 359 Db 241501 AATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAAAAAATCGATTGGCGTGCCTTC 241560	
gene	360 AGTCCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAGAGATGCAAG 419 Db 241561 AGTCCTTGTAGGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAGAGATGCAAG 241620	
gene	420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAATGATGATTCAG 479 Db 241621 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAATGATGATTCAG 241680	
gene	480 TAGCGTTAGTGTCTCCACTTAAACAGACGACACAAACGTTGAGTGATGATCTAAACATCGTC 539 Db 241681 TAGCGTTAGTGTCTCCACTTAAACAGACGACACAAACGTTGAGTGATGATCTAAACATCGTC 241740	
gene	540 AATCACTAATAATGCGCAACAGAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 599 Db 241741 AATCACTAATAATGCGCAACAGAGTGTGGCGCAAAATCCAGCACAACAGGAAACGACACA 241800	
gene	600 ATCACTATCAACAATGCACTACGGAAGAAACCCCGTAACTCGTGAAGCTACTACTAC 659 Db 241801 ATCACTATCAACAATGCACTACGGAAGAAACCCCGTAACTCGTGAAGCTACTACTAC 241857	
gene	660 GACAAAGAAATCAAGCTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 719 Db 241858 GACAAAGAAATCAAGCTAATACACCGCAACAACTCAATCAAGCAATCAAAATCGGAGGA 241917	
gene	720 ATTAGTGAATCAACAAGTAAATGAACGACTTTTAAATGATGATGATGATGATGATGAT 779 Db 241918 ATTAGTGAATCAACAAGTAAATGAACGACTTTTAAATGATGATGATGATGATGATGAT 241977	
gene	780 AAATTCACCTCAAAATTTCTCAAAATGCGGAAAATGTTTCAACACGCAAGATCTTCAAC 839 Db 241978 AAATTCACCTCAAAATTTCTCAAAATGCGGAAAATGTTTCAACACGCAAGATCTTCAAC 242037	
gene	840 TGAAGCAACACCTTCAAAACAAATGAATCAGTCCACAGATGACAGATCAAGTAATAAGA 899 Db 242038 TGAAGCAACACCTTCAAAACAAATGAATCAGTCCACAGATGACAGATCAAGTAATAAGA 242097	
gene	900 TGTAGTTAATCAACGGTGTAAATCAAGTGGCCCTAGTAATGAGAGCAATTTAGTTAGCGGC 959 Db 242098 TGTAGTTAATCAACGGTGTAAATCAAGTGGCCCTAGTAATGAGAGCAATTTAGTTAGCGGC 242157	
gene	960 AGTAGTGCAGATGCACCGGACCTGGCAGACAGATATTACGAATCAGTTGACGAATGTGAC 1019 Db 242158 AGTAGTGCAGATGCACCGGACCTGGCAGACAGATATTACGAATCAGTTGACGAATGTGAC 242217	
gene	1020 AGTTGGTATTGACTCTGGTACGACTGTGTATCCGACCAAGAGGTTATGTCAAATGAA 1079 Db 242217 AGTTGGTATTGACTCTGGTACGACTGTGTATCCGACCAAGAGGTTATGTCAAATGAA 242277	

Query Match 92.8%; Score 3245; DB 1; Length 290150;

Db	242218	AGTTGGTATTGACTCTGGAGATACAGTCTTATCCGACCAAGCAGCGCTATGTCAAACCTGAA	242277
Qy	1080	TTATGGTTTTTCAGTGCCTTAATTCCTGCTTTAAAGGTGACACATTCACAAATAAAGTGTACC	1139
Db	242278	TTATGGTTTTTCAGTGCCTTAATTCCTGCTTTAAAGGTGACACATTCACAAATAAAGTGTACC	242337
Qy	1140	TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCACCAATTAATGCTGG	1199
Db	242338	TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCCTTCAATTAATGSCCGG	242397
Qy	1200	AGATCAAGTATTGGCAATGGTGTAAATCGATAGTGTAAATGTTTATTTATACATTTAC	1259
Db	242398	AGATCAAGTATTGGCAATGGTGTAAATCGATAGTGTAAATGTTTATTTATACATTTAC	242457
Qy	1260	AGATCAATGTAAATPACTAAAGATGATGTAAAGCAACTTTTGACCATGCCGCTTATATTGA	1319
Db	242458	AGATCAATGTAAATPACTAAAGATGATGTAAAGCAACTTTTGACCATGCCGCTTATATTGA	242517
Qy	1320	CCCTGAAATGTTTAAAGACACAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAAC	1379
Db	242518	CCCTGAAATGTTTAAAGACACAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAAC	242577
Qy	1380	AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTTTATAACTTTATCTAT	1439
Db	242578	AGCAAAACAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTTTATAACTTTATCTAT	242637
Qy	1440	TAAAGGTACAAATTGACCAATTCGATTAACAAATAATACGATTCGTCAGACAATTTATGT	1499
Db	242638	TAAAGGTACAAATTGACCAATTCGATTAACAAATAATACGATTCGTCAGACAATTTATGT	242697
Qy	1500	CAATCCAGTGGAGATACGTTATTCGCCGGTTTTTAAACAGGTAATTTTAAACCAATAAC	1559
Db	242698	CAATCCAGTGGAGATACGTTATTCGCCGGTTTTTAAACAGGTAATTTTAAACCAATAAC	242757
Qy	1560	GGATAGTAATGCATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAA	1619
Db	242758	GGATAGTAATGCATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAA	242817
Qy	1620	TGCAGCTGATTTTACTGAAAGTTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACTAA	1679
Db	242818	TGCAGCTGATTTTACTGAAAGTTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACTAA	242877
Qy	1680	TAGTGTGAATTTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATAGCCCTGA	1739
Db	242878	TAGTGTGAATTTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATAGCCCTGA	242937
Qy	1740	TGATCAAAATACAAACACCGTATATAGTAGTTTAAATGGTCATATTTGATCCGAATAGCAA	1799
Db	242938	TGATCAAAATACAAACACCGTATATAGTAGTTTAAATGGTCATATTTGATCCGAATAGCAA	242997
Qy	1800	AGGTGATTTAGCTTTAGCTTCAACTTTATATGGGTATAACTCGAATATAATTTGCGCGTC	1859
Db	242998	AGGTGATTTAGCTTTAGCTTCAACTTTATATGGGTATAACTCGAATATAATTTGCGCGTC	243057
Qy	1860	TATGTATGGACAAACAGTAGCATTTAATAACGGATCAGGTTCTGTGACGGTATCGA	1919
Db	243058	TATGTATGGACAAACAGTAGCATTTAATAACGGATCAGGTTCTGTGACGGTATCGA	243117
Qy	1920	TAAACCAAGTTGTTCTGTAACCAACCTGATGAGCCTGGTGAATTTGAACCAATTCAGAGGA	1979
Db	243118	TAAACCAAGTTGTTCTGTAACCAACCTGATGAGCCTGGTGAATTTGAACCAATTCAGAGGA	243177
Qy	1980	TTTCAGATTTCAACCGAGTTTCAGATTTCTGGCAGCGATTCTTAATTCAGATAGCGGTTTACA	2039
Db	243178	TTTCAGATTTCAACCGAGTTTCAGATTTCTGGCAGCGATTCTTAATTCAGATAGCGGTTTACA	243237
Qy	2040	TTCCGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGTTCAGATTCAGC	2099
Db	243238	TTCCGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGTTCAGATTCAGC	243297
Qy	2100	AAGTGATTCAGACTCAGCGAGTGTTCAGATTCAGCAAGCGATTCGGAATTCAGCGAGCGA	2159
Db	243298	AAGTGATTCAGACTCAGCGAGTGTTCAGATTCAGCAAGCGATTCGGAATTCAGCGAGCGA	243357
Qy	2160	TTCCGACTCAGCAATGACTCCGATTTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGA	2219
Db	243358	TTTCAGATTCAGCGAGCGATTTCAGATTCAGCGAGTGAATTCAGATTCAGCGAGTGAATTCAGA	243417
Qy	2220	T-----TCCGA	2225
Db	243418	TTTCAGCAAGCGATTCCGACTCAGACAAATGACTCCGATTTCAGATAGCGATTCTGACTCAGA	243477
Qy	2226	CAGTGAATTCAGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGATTCAGATAGCGA	2285
Db	243478	CAGTGAATTCAGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGATTCAGATAGCGA	243537
Qy	2286	TTTCAGATTCAGATAGCGATTTCAGATTCAGACAGTGAATTCGGAATTCAGACAGCGATTCTGA	2345
Db	243538	TTTCAGATTCAGATAGCGATTTCAGATTCAGACAGTGAATTCGGAATTCAGACAGCGATTCTGA	243597
Qy	2346	CTCCGACAGTGAATTCGGAATTCAGACAGCGATTTCAGATTCGGAATTCGGAATTCGGAATTCGGA	2405
Db	243598	CTCCGACAGTGAATTCGGAATTCAGACAGCGATTTCAGATTCGGAATTCGGAATTCGGAATTCGGA	243657
Qy	2406	TAGCGATTCGGAATTCAGATAGCGACTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGA	2465
Db	243658	TAGCGATTCGGAATTCAGATAGCGACTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGA	243717
Qy	2466	TTTCAGATTCAGATAGCGATTTCAGATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	2525
Db	243718	TTTCAGATTCAGATAGCGATTTCAGATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	243777
Qy	2526	TTTCAGATAGCGATTTCAGATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	2585
Db	243778	TTTCAGATAGCGATTTCAGATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	243837
Qy	2586	CAGTGAATTCGGAATTCAGCGAGTGAATTCGGAATTCAGATAGTGAATTCGGAATTCGGAATTCGGA	2645
Db	243838	CAGTGAATTCGGAATTCAGCGAGTGAATTCGGAATTCAGATAGTGAATTCGGAATTCGGAATTCGGA	243897
Qy	2646	CTCCGATTCAGATAGCGACTTCAGATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	2705
Db	243898	CTCCGATTCAGATAGCGACTTCAGATTCGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	243957
Qy	2706	CTTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGAC	2765
Db	243958	CTTCAGATAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGACAGCGATTTCAGATTCAGAC	244017
Qy	2766	CAGCGACTCAGACAGTGAATTCAGATTCAGATAGTGAATTCGGAATTCGGAATTCGGAATTCGGA	2825
Db	244018	CAGC-----GACTTCAGATTCAGATAGTGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	244065
Qy	2826	CTCAGGTAGTGAATTCAGATTCAGATAGTGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	2885
Db	244066	CTCAGGTAGTGAATTCAGATTCAGATAGTGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	244125
Qy	2886	TTCCGAGTCAAGTTCTAAACAAATTAATAGTGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	2945
Db	244126	TTCCGAGTCAAGTTCTAAACAAATTAATAGTGAATTCGGAATTCGGAATTCGGAATTCGGAATTCGGA	244185
Qy	2946	TGCTTCTAAATAAATGAGGTGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAAT	3005
Db	244186	TGCTTCTAAATAAATGAGGTGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAATAGTGAAT	244245
Qy	3006	AGATGAAGCAAAATACGTCACCTAAATTTGGGATTTATAGCATCAATAGTGTCTATTTACTTACT	3065
Db	244246	AGATGAAGCAAAATACGTCACCTAAATTTGGGATTTATAGCATCAATAGTGTCTATTTACTTACT	244305
Qy	3066	TTTCAGAGCAAAATAAAGATAAAGATAAAGATAAAGATAAAGATAAAGATAAAGATAAAGATAAAG	3125
Db	244306	TTTCAGAGCAAAATAAAGATAAAGATAAAGATAAAGATAAAGATAAAGATAAAGATAAAGATAAAG	244365
Qy	3126	ATATGATTCATGAAGAGCCACCTTAAAGGTGCTCTTTTACTTGGATTTTCCAAATAT	3185
Db	244366	ATATGATTCATGAAGAGCCACCTTAAAGGTGCTCTTTTACTTGGATTTTCCAAATAT	244425


```
/product="MnhD homologue"
/protein_id="BAB56787.1"
/db_xref="GI:14246394"
/translation="MLSNLLILPMLLPFLCALILVFLKNDRIISKYLYLGTWITITII
SLMLLIYVORHPITLDFGWSAPGQIFGLDSLSIMVWTASVITILIMAYGFRGE
HKARYHLVPELIFSLVGVIGFSLDFPNLYVMPFIMLLASVFLITLQSQVEOLRAA
IIVVNLIGSWLFLGLGLYKTVGTILFNSHIANRLNDMGDNRTVMISLIFLVAFS
AKAALVFLMPLPKAYAVNLTAAALFAALMTKVGAYALIRFTLLFDORHLLIPELLA
TMAAITMVGAIYAYDKIKKIAAYQVILSIGFILGLGNTFAGINGALFYLVNDLI
VKTLFLFISLVITGVYQVYGLNGLAKEPELFGVAFIIMFAIGVPPFSGPFGK
VLIFOGALONGNYIGLAMITISLIAMYSLEFRIFPFYMGDKGSEVNFKIPLYRKR
ILSLVVVIAIGIAAPVVLNVTSATELNTSDQLYQKLVNPHLKGED"
5327..5809
/locus_tag="SAV0626"
5327..5809
/locus_tag="SAV0626"
/codon_start=1
/transl_table=11
/product="Na+ antiporter"
/protein_id="BAB56788.1"
/db_xref="GI:14246395"
/translation="WNQIVLINIIIFLWVLFQDEHFKFSTFFSGYLIGLIVYILHR
FSDDFYVRKIWAIFKFLGVLYQILTSSITINILFKTMDNPFLLSYETRLTSDW
AITEFLIILITPGSVIRISQDSKKFHSIDVSEKEKDSLRSIKHYEDLILEVSR
"
5806..6108
/locus_tag="SAV0627"
5806..6108
/locus_tag="SAV0627"
/codon_start=1
/transl_table=11
/product="similar to Na+ antiporter"
/protein_id="BAB56789.1"
/db_xref="GI:14246396"
/translation="MIQTTHIMIISSLIIFGIALIICLPRIKGPPTADRVVTFDIT
SAVMSIVGLSVLMTGVFLDSIMLIAISFVSSVSIISRFIGGHVFNNGNKNRL"
6083..6520
/locus_tag="SAV0628"
6083..6520
/locus_tag="SAV0628"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB56790.1"
/db_xref="GI:14246397"
/translation="MBITKEIFSLIAAVMLLGSFIALISAIGIVKFQDFVFLSHAAT
KSTLSVLTLIGLVIFVNTGPFVSRLSLVFINLTPVGMHLVARAAYRNGAYM
YRKNDATHASILLSSNEQNSTEALQLRAKKREHRKKWYQND"
6860..8902
/locus_tag="SAV0629"
6860..8902
/locus_tag="SAV0629"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB56791.1"
/db_xref="GI:14246398"
/translation="MEIFETPILIFIAVVLISFVHTFIPKVPFLAFIQIFLGMILFIFTP
IPVQNFDSFLFMTWIAPIILFVEGVNSRVHLRKYIKFPMVMALGLVITTVIGLGF
IHWIWPDLPGAFAIAILCPTDAVAQAITKGKLPKGMAMTILEGESILINDAAGII
SFKIAGVLTGTASLDVQFLIASIGVAVGILLGMALVREFLIRMRGYNINM
FTIQLTPTFVYTLIAELFHASGIIAAVAGLVHGFERDRIMQVRLQMSYHTWNI
LGVYLAISPFQKMTKDDNPTEKPPKPHNLIFLIGITIVHALAVLFRVWYVLY
PFYVLAISPFQKMTKDDNPTEKPPKPHNLIMTLICGVHTIISLAIALTLPIYEL
AGHAFYRNDLLFIASGMVILISVAQVLLPLTKPKPTVIGNMSPKVARIVILEQ
VIDINQKSTFETSKYGNVIEKHDKLAFILKTEKODENSEKELERLQKIAFNVEKTI
LESVDEQIITNSVLENTVQVRSALRMIIVLRGALLKRVQTRVNSAS
LSVTDNLMEINKLKVHYNVSRSLSKETTKDNTLEIGVMVCDGLVIMRIENLTSPNFF
NSASEDTTKIKLNAALRQRRLIRELIDTDEVSEGTALKLREAINYDEMWIVDSMT"
complement(9010..10329)
/gene="tnp"
/locus_tag="SAV0630"
complement(9010..10329)
```

```
/gene="tnp"
/locus_tag="SAV0630"
/codon_start=1
/transl_table=11
/product="transposase for IS1181"
/protein_id="BAB56792.1"
/db_xref="GI:14246399"
/translation="WCNDTLELLRIKDENIKYINQEIIVIIKKGKATVNVAVLYTKPS
"
Query Match 84.6%; Score 2960.6; DB 1; Length 348477;
Best Local Similarity 90.7%; Pred. No. 0; Mismatches 319; Indels 7; Gaps 2;
Matches 3179; Conservative 0;

QY 1 GGTACACATAAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTACATTA 60
DB 193322 GGTACACATAAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTACATTA 193381

QY 61 GGTAGAGTTTCATTAATAATAAAAAATGTTTGCAATCAATCGTACGTTGTCGTTGTA 120
DB 193382 GGTAGAGTTTCATTAATAATAAAAAATGTTTGCAATCAATCGTACGTTGTCGTTGTA 193441

QY 121 ATTCTTAAATAGCAATAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA 180
DB 193442 ATTCTTAAATAGCAATAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA 193501

QY 181 TATCGATACAAAATTAAATGCTATAATGCAATTTTAGTGTATTAATTCATTAAACAGAGATT 240
DB 193502 TATCGATACAAAATTAAATGCTATAATGCAATTTTAGTGTATTAATTCATTAAACAGAGATT 193561

QY 241 AAATATATC-TTAAAGGGTATATAGTTAATAATAAATGATTTTAAAAAGGGGATTA 299
DB 193562 AAATATATCTTTAAAGGGTATATAGTTAATAATAAATGATTTTAAAAAGGGGATTA 193621

QY 300 AATGAATATGAAGAAAAGAAAACACGCAATTCGGAATAAATCGATTGGCGTGGCTTC 359
DB 193622 AATGAATATGAAGAAAAGAAAACACGCAATTCGGAATAAATCGATTGGCGTGGCTTC 193681

QY 360 AGTGTCTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAAGAACAGATGCAAG 419
DB 193682 AGTGTCTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAAGAACAGATGCAAG 193741

QY 420 TGAATAATAGTTTACGCAATCTGATCGGCAAGTACGAAAGCAAAAGTAATGATTCAG 479
DB 193742 TGAATAATAGTTTACGCAATCTGATCGGCAAGTACGAAAGCAAAAGTAATGATTCAG 193801

QY 480 TAGCGTTAGTGTGACCTTAAACACAGACACAAACGTCAGTGATATACTAAAAACATCGTC 539
DB 193802 TAGCGTTAGTGTGACCTTAAACACAGACACAAACGTCAGTGATATACTAAAAACATCGTC 193861

QY 540 AAACACATTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACAAACAGAAACGACACA 599
DB 193862 AAACACATTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACAAACAGAAACGACACA 193921

QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTAC 659
DB 193922 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTAC 193981

QY 660 GACAAAGAAATCAAGCTAATACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGGA 719
DB 193982 GACAAAGAAATCAAGCTAATACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGGA 194041

QY 720 ATTAGTGAATCAAAACAAGTAATGAAACGACTTTTAAATGATACTAATACAGTATCATCTGT 779
DB 194042 ATTAGTGAATCAAAACAAGTAATGAAACGACTTTTAAATGATACTAATACAGTATCATCTGT 194101

QY 780 AAATTCACCTCAAAATTTCTACAAATCGGAAAAATGTTTCAACACGCAAGTACTTCAAC 839
DB 194102 AAATTCACCTCAAAATTTCTACAAATCGGAAAAATGTTTCAACACGCAAGTACTTCAAC 194161

QY 840 TGAAGCAACACCTTCAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 899
DB 194162 TGAAGCAACACCTTCAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 194221
```

QY	900	TGTAGTTAATCAAGCGTTAAATACAAAGTGGCGCTAGAAATGAGAGCAATTTAGTTAGCGGC	959	Db	195302	TTCCAGATTTCTGACCCAGGTTTCAGATTTCTGGCAGGATTTCTAATTCAGATAGCGGTTTCAGA	195361
Db	194222	TGTAGTTAGTCAAGCGGTTAAATCCAAAGTACGCGCTAGAAATGAGAGCAATTTAGTTAGCGGC	194281	QY	2040	TTCCGGGTAGTGAATTCACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGC	2099
QY	960	AGTAGTGCAGATGACACCGGAGCTGGCACAGATATTACGAATTCAGATTCAGCGAATGTGAC	1019	Db	195362	TTCTGGCAGTGAATTTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGC	195421
Db	194282	AGTAGTGCAGATGACACCGGAGCTGGCACAGATATTACGAATTCAGATTCAGCGAATGTGAA	194341	QY	2100	AAAGTGAATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGCGATTCGGACTTCAGCGAGCGA	2159
QY	1020	AGTTGGTATGTACTCTGGTACGACTGTGTATCCGCACCAAGCAGGTTATGTCAAACTGAA	1079	Db	195422	AAAGTGAATTCAGACTCAGCGAGTGAATTCAGATTCAGCGAGCGATTCAGATTCAGCAAGTGA	195481
Db	194342	AGTTACTATGTACTCTGGTACGACTGTGTATCCGCACCAAGCAGGTTATGTCAAACTGAA	194401	QY	2160	TTCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGA	2219
QY	1080	TTATGCTTTTTCAGTGCCTAAATCTGCTGTTAAAGGTGACATTCMAAATAACTGTACC	1139	Db	195482	TTCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGA	195541
Db	194402	TTATGCTTTTTCAGTGCCTAAATCTGCTGTTAAAGGTGACATTCMAAATAACTGTACC	194461	QY	2220	TTCCGACTCAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	2279
QY	1140	TAAAGAAATTAACCTTAAATGGTGTAACTCAACTGCTAAAGTGCACCAATTAATGCTGG	1199	Db	195542	TTCCGACTCAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	195601
Db	194462	TAAAGAAATTAACCTTAAATGGTGTAACTCAACTGCTAAAGTGCACCAATTAATGCTGG	194521	QY	2280	TAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	2339
QY	1200	AGATCAAGTATTTGGCAATGGTGTAAATCGATAGTGGTAAATGTTATTTATACATTTAC	1259	Db	195602	GAGTGAATTCGACTCAGCGAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	195661
Db	194522	AGATCAAGTATTTGGCAATGGTGTAAATCGATAGTGGTAAATGTTATTTATACATTTAC	194581	QY	2340	TTCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGACAGTGAATTCGGA	2399
QY	1260	AGACTATCTAAATCTAAAGATGATGTAAGCAACTTTGACCAATTTGACCGCGCTTATATGA	1319	Db	195662	TTCCGACTCAGATAGCGACTCAGATTCAGACAGCGATTTCTGACTCAGACAGCGATTCCTGA	195721
Db	194582	AGACTATCTAAATCTAAAGATGATGTAAGCAACTTTGACCAATTTGACCGCGCTTATATGA	194641	QY	2400	CTCAGATAGCGATTTCCGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	2459
QY	1320	CCCTGAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAC	1379	Db	195722	CTCAGATAGCGATTTCCGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	195781
Db	194642	CCCTGAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAC	194701	QY	2460	CAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	2519
QY	1380	AGCAAAACAAACAGTATTTAGTATGATTAATGAAATATATCGTAAGTTTATATACTTATCAT	1439	Db	195782	TAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	195841
Db	194702	TGCTAGTAAAGACAGTATTTAGTATGATTAATGAAATATATCGTAAGTTTATATACTTATCAT	194761	QY	2520	CTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGA	2579
QY	1440	TAAAGGTACAAATTCGACCAATTCGATAAACAATTAATACGTATCGTACAGCAATTTATGT	1499	Db	195842	CTCAGACTCAGACAGCGATTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	195901
Db	194762	TAAAGGTACAAATTCGACCAATTCGATAAACAATTAATACGTATCGTACAGCAATTTATGT	194821	QY	2580	CTCAGACTCAGACAGCGATTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	2639
QY	1500	CAATCCAAAGTGGAGATAACGTTATTGCGCGGTTTAAACAGGTAATTTAAACCAAAATAC	1559	Db	195902	TTCCGATAGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	195961
Db	194822	CAATCCAAAGTGGAGATAACGTTATTGCGCGGTTTAAACAGGTAATTTAAACCAAAATAC	194881	QY	2640	CAGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	2699
QY	1560	GGATAGTAAATGATATAGTACAGCAAAATACAAATTAATTAAGTATATAAGTAGATAA	1619	Db	195962	TAGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	196021
Db	194882	AAAGAGTAAATGATATAGTACAGCAAAATACAAATTAATTAAGTATATAAGTAGATAA	194941	QY	2700	TTCCGACTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	2759
QY	1620	TGACAGTGAATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTGAGATGTCACTAA	1679	Db	196022	TTCCGACTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	196081
Db	194942	TGCTAATGATTTATCTGAAAGTTACTTTGTGAATCCCTAGCGATTTTGAAGATGTAACCTAA	195001	QY	2760	TTCCGACTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	2813
QY	1680	TAGTGAATTTATCAATTTCCAAATCCAAATCAATATAAAGTAGAGTTTAAATCGCCTGA	1739	Db	196082	TTCCGACTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	196141
Db	195002	TCAAGTTAGAATTTCAATTTCCAAATCCAAATCAATATAAAGTAGAGTTTAAATCGCCTGA	195061	QY	2814	GAGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	2873
QY	1740	TGATCAATTTACACACCGTATATAGTATGTTTAAATGGTCAATATTCGATTCGATTCGAA	1799	Db	196142	AAAGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	196201
Db	195062	TGACCAATTTACACACCGTATATAGTATGTTTAAATGGTCAATATTCGATTCGATTCGAA	195121	QY	2874	TTCAATAGCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	2933
QY	1800	AGGTGAATTTAGCTTTCACTTTTATATGGGTAAATCTCGATATAATTTGGCGCTC	1859	Db	196202	TTCAATAGCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	196261
Db	195122	AGGTGAATTTAGCTTTCACTTTTATATGGGTAAATCTCGATATAATTTGGCGCTC	195181	QY	2934	AAATGGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	2993
QY	1860	TATGTATGGGACCAAGTAGCATTTAATTAACGGATCAGGTTTCTGGTACCGGTATCGA	1919	Db	196262	AAATGGTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCAGA	196321
Db	195182	TATGTATGGGACCAAGTAGCATTTAATTAACGGATCAGGTTTCTGGTACCGGTATCGA	195241	QY	2994	TACAGGTTTCTGAAGATCAAGCAATACGTCACCTAAATTTGGGGATTTATAGCATCAATAGG	3053
QY	1920	TAAACAGTGTTCCTGGAACAACCTGATGAGCGCTGGTGAATTTGAACCAATTCAGAGGA	1979	Db	196322	TACAGGTTTCTGAAGATCAAGCAATACGTCACCTAAATTTGGGGATTTATAGCATCAATAGG	196381
Db	195242	TAAACAGTGTTCCTGGAACAACCTGATGAGCGCTGGTGAATTTGAACCAATTCAGAGGA	195301	QY	3054	TTCAATTTACTCTTTTTCAGAGGAAAAAAGAAATAAGATAAGAAATAAGATAAGATATATGAT	3113
QY	1980	TTCCAGATTTCTGACCCAGGTTTCAGATTTCTGCGAGCGATTTCTAATTCAGATAGCGGTTTCA	2039				


```
gene
CDS
DDPKGCGSGSLMNLILQSNFNRHRAIVDKGVLEKACSTLLTTFFKNLRANKKSTN"
4329..5198
/gene="SA0517"
4329..5198
/gene="SA0517"
/note="ORFID:SA0517"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB41748.1"
/db_xref="GI:13700451"
/translacion="MKLIATDMGTLNAAHEISQPINDAIKYAEQOGITVVIATGR
AFYERAAQAVATDLTVPICLNGAEVROETFMVSTSHLSLHKHINVLKAGIYY
QYYSRIYSDPQDLDIYIDIAERAGHAENVERIKSGIQRINDIGTKVDNDYDAI
ENIPELIMKLAFDNELEKIDKASKILAEANFNLAISSSRGNIEITHSDAKGTALE
TIAERLIGLEMDQVAIGNDLMSLEKVGYPVAMENGABEEVKIAKIVTDTNENSGV
GKAIWKILREKQV"
5220..5786
/gene="SA0518"
5220..5786
/gene="SA0518"
/note="ORFID:SA0518"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB41749.1"
/db_xref="GI:13700452"
/translacion="MKGLIILIGSAQVNSHTSALARYLPEHPKTHDIEAEIPLAEKPL
NQLDFSGTPTSIDIKOMKDLKEMAAADPLILGTNYHGSYGLKALNDHLNMDY
FKMPVGLIGSGGIVSESEPLSHLRVIRVSLIGIAVPIQIATHSDSDFAKNEDGGYYLN
DSEFLQRLARLEVDQIVSFVNNSPYEHLK"
6215..9076
/gene="sdrC"
6215..9076
/gene="sdrC"
/note="ORFID:SA0519"
/codon_start=1
/transl_table=11
/product="Ser-Asp rich fibrinogen-binding, bone
sialoprotein-binding protein"
/protein_id="BAB41750.1"
/db_xref="GI:13700453"
/translacion="MNNKKTATNRKGMIPNRLNFKSRKYSVGTASILVGTTLIFGSL
GHEKAAHNGELNMQSKNETTAPSENKTTIEKVDROLKQNTOTATADQPKVMSDSA
TVKETSNNQPNATASQSTQTSNVTTNDKSTTYSDNETKSLNLTQAKNVSTPTKT
TTIKRALNRMAVNTAAPQOGTNDVKVFTNIDIAIDKHVNKTTGNTFEFATSSD
VLKLANYTIDDSVKEGDTFTFKYGOVFRPGSVRLPSQTNQNLNAQGNIIAKGLYDSK
TNTTYFTNVVDQYTNVSGSFEQVAFKRENATTDKTXMEVTLGNDIYSKDVIVD
YGNKGOQLISSINYINNELSRNMTVYNQPKTYTKETFPVITNLITGKPNPAKPK
IYEVTDNQFVDSFTPTSKLDVITGQFDVIYSDNKTATVDLLNGSSDRQYIIQQ
VAYPDSNTDNGKIDITYLETQNGKSSNSYSNVNGSSTANGQPKYKYNLGDYVWEDTN
KDGQDANEGKIKGVYVILKDSNGKELDRITTDENGKYQFTGLNSYYSVEFSPAGY
TPTTANAGTDDVDSDGLTGTGVIKADNMNLDLSDGYKFTPKYSLGDYVWVDSNKGQK
DSTEKIGIKGVKVLQNEKGVIGTETEDNGKVFEDNLDGKVKVFEKPAGLTQGTG
NTTEDDQADQGVVDVITLDDHDFTLNGYYEETSDDSDSDSDSDSDSDSDSDSDSD
DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
DSDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
ALPETGSENNNSNNGTLFGGLFAALGSLILFGRKKQNK"
9443..13600
/gene="sdrD"
9443..13600
/gene="sdrD"
/note="ORFID:SA0520"
/codon_start=1
/transl_table=11
/product="Ser-Asp rich fibrinogen-binding, bone
sialoprotein-binding protein"
/protein_id="BAB41751.1"
/db_xref="GI:13700454"
/translacion="MLNRENKTAITRGMVSNRLNFKFSIKYTVGTASILVGTTLIFG
LGNOEAKAESTNKLNEATTSASDNQSSDKVDMQOLNQEDNTKDNOKEMWSSQNE
TTSNGNKSIEKESVQSTGNGKVEVSTAKSDEQASPKSTNEDLNTKQITISNQELQFDL
```

```
Query Match 81.6%; Score 2853.4; DB 1; Length 299050;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 3232; Conservative 0; Mismatches 266; Indels 169; Gaps 5;

QY 1 GGTACCAATAATTACACATCTGCTTTTGAAGAAATATGATTTCAAGCTAGGATTACATTA 60
DB 249193 GGTACCAATAATTACACATCTGCTTTTGAAGAAATATGATTTTAAAGCTAGGATTACATTA 249242
QY 61 GGTAGAGTTCATATTAATAATAAAAAATGTTTGAATCAAAATCGTACCTTGTGCTTGTGA 120
DB 249243 GGTAGAGTTCATATTAATAATAAAAAATGTTTGAATCAAAATCGTACCTTGTGCTTGTGA 249302
QY 121 ATTCTTAAAAATAGCAATAATAATAATGTTTGTAGTAAAGTATTATTGTGGAATAATAAAA 180
DB 249303 ATTCTTAAAAATAGCAATAATAATAATGTTTGTAGTAAAGTATTATTGTGGAATAATAAAA 249362
QY 181 TATCGATACAAATTAATGCTATTAATGCAATTTTAGTGTATAAATCCATTACAGAGATT 240
DB 249363 TATCGATACAAATTAATGCTATTAATGCAATTTTAGTGTATAAATCCATTACAGAGATT 249422
QY 241 AAATATATC-TTAAAGGGTATATAGTTAATAATAAAATGACTTTTTAAAAAGAGGGAATAA 299
DB 249423 AAATATATCTTTAAAGGGTATATAGTTAATAATAAAATGACTTTTTAAAAAGAGGGAATAA 249482
QY 300 AATCAATATGAGAAAAAAGAAAAACAGCAATTCGGAAAAAATCGATTGCGGTGGCTTC 359
DB 249483 AATCAATATGAGAAAAAAGAAAAACAGCAATTCGGAAAAAATCGATTGCGGTGGCTTC 249542
QY 360 AGTCTTGTAGTAGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAAAGAAAGCAGATGCAAG 419
DB 249543 AGTCTTGTAGTAGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAAAGAAAGCAGATGCAAG 249602
QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAAAGTAAATGATTCAAG 479
DB 249603 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAAAGTAAATGATTCAAG 249662
QY 480 TAGCGTTAGTGTGCTCACTTAAACAGACAGACAAACGTCGAGTGATATCTTAAACATCGTC 539
DB 249663 TAGCGTTAGTGTGCTCACTTAAACAGACAGACAAACGTCGAGTGATATCTTAAACATCGTC 249722
QY 540 AAACACTATAATAATGCGGAAACGAGTGTGGCGCAAAATCCAGCACAACAGAGAAACGACACA 599
DB 249723 AAACACTATAATAATGCGGAAACGAGTGTGGCGCAAAATCCAGCACAACAGAGAAACGACACA 249782
QY 600 ATCATATCAACAAATGCAACTACGGAAGAAACCGCGGTAACTGGTGAAGCTACTACTAC 659
DB 249783 ATCATATCAACAAATGCAACTACGGAAGAAACCGCGGTAACTGGTGAAGCTACTACTAC 249842
QY 660 GACACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 719
DB 249843 GACACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGA 249902
QY 720 ATTAGTGAATCAACAAAGTAAATGAAACGACTTTTAAATGACTACTAATACAGTATCATCTGT 779
DB 249903 ATTAGTGAATCAACAAAGTAAATGAAACGACTTTTAAATGACTACTAATACAGTATCATCTGT 249962
QY 780 AATTTCACCTCAAAATTTCAAAATCGGGAATAATGTTTCAACCAACCGCAAGATCTTCAAC 839
```

Db	249963	AAATTCACTCCAAAATTTCTCAAATGCGGAAAAATGGTTCACAAAGCAGATACACTTCAAC	250022
Qy	840	TGAAGCAACACCTTTCAAAACAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAAGA	899
Db	250023	TGAAGCAACACCTTTCAAAACAATGNACAGCTCCACAGAAATACAGATGCAAGTAATAAAGA	250082
Qy	900	TGTAGTTAAATCAAGCGGTTAAATCAAGTGCGCTTAGAATGAGAGCAATTTAGTTTAGCGGC	959
Db	250083	TGTAGTTAGTCAAGCGGTTAAATCAAGTAGCGCTTAGAATGAGAGCAATTTAGTTTAGCGGC	250142
Qy	960	AGTAGCTGAGATGCCACCGCAGCTGGCACAGATATTTACGAATCAGTTGACGAAATGTGAC	1019
Db	250143	AGTAGCTGCGATGCGACCGGAGCTGGCACAGATATTTACGAATCAGTTGACGAAATGTGAA	250202
Qy	1020	AGTTGGTATTGACTCTGCTGACGACTGTGTATCCGCACCAAGCAGGTTATGTCACAACTGAA	1079
Db	250203	AGTTTACTATTGACTCTGCTGACGACTGTGTATCCGCACCAAGCAGGTTATGTCACAACTGAA	250262
Qy	1080	TTATGGTTTTTCAGTGCGCTAATTTCTGCTGTTTAAAGGTGCACATTTCAAAATAACTGTACC	1139
Db	250263	TTATGGTTTTTCAGTGCGCTAATTTCTGCTGTTTAAAGGTGCACATTTCAAAATAACTGTACC	250322
Qy	1140	TAAAGAAATTAACCTTAAATGGTGTAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGG	1199
Db	250323	TAAAGAAATTAACCTTAAATGGTGTAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGG	250382
Qy	1200	AGATCAAGTATTGGCAAAATGCTGAATCGATGATGATGTAATGTTATTTATACATTTAC	1259
Db	250383	AGATCAAGTATTGGCAAAATGCTGAATCGATGATGATGTAATGTTATTTATACATTTAC	250442
Qy	1260	AGACTATGTAATAATCTAAAGATGATGTAAGAACCACTTTGHACCATGGCCGCTTATTTGA	1319
Db	250443	AGACTATGCTGATAATAAAGAAAAATGTAAACAGCTAAATATTACTATGCGAGCTTATTTGA	250502
Qy	1320	CCCTGAAAATGTTAAAAAGACAGGTAATGTGCATTTGGCTACTCGCATPAGTGTACAAAC	1379
Db	250503	CCCTGAAAATGTTAAAAAGACAGGTAATGTGCATTTGGCTACTCGCATPAGGAAACCAATAC	250562
Qy	1380	AGCAAAACAAAACAGTATTAGTATGATTATGAAAATATCGTAAAGTTTTTATAACTTATCTAT	1439
Db	250563	TGCTAGTAAAGACAGTATTAAATCGACTATGAGAAATATGCAAAATTTCCATAATTTATCAAT	250622
Qy	1440	TAAAGGTACAAATTGACCAATTCGATPAAAACAAAATAATACGTATCGTCAGACAAATTTATGT	1499
Db	250623	TAAAGGTACGATTTGATCAAAATCGATAAAACAAAATAATACGTATCGCCAAAACAAATTTATGT	250682
Qy	1500	CMAATCCAAGTGAGATTAACGTTATTTGGCCGGTTTTTAACAGGTAAATTTAAACCAAAATAC	1559
Db	250683	CAATCCAAGCGGAGATAACGTTGTGTACCTGCCCTTAACAGGTAAATTTAAATTCCTAATAC	250742
Qy	1560	GGATAGTAATGCATTTAAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAGTAGATAAA	1619
Db	250743	AAAGAGTAATGCGTTAAATAGATGCAAAAACAACTGATATTAAAGTTTATAGATCGATAAA	250802
Qy	1620	TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCACCTAA	1679
Db	250803	TGCTAATGATTTATCTGAAAGTTATTTATGTGAATCCTAGCGATTTTGAAGATGTAACTAA	250862
Qy	1680	TAGTGTGAATATTAATCTCCAAAATCCAAATCAATATAAGTAGAGTATTAAATACCCCTGA	1739
Db	250863	TCAAGTTAGAAATTTCAATTTCCAAAATGCTAATCAATACAAAGTAGAATTTCTTACGACGA	250922
Qy	1740	TGATCAAAATTAACAACCGTATATAGTGTGTTAATGCTCATATTGATTCGGAATAGCAAA	1799
Db	250923	TGACCAAAATTAACAACCGTATATAGTGTGTTAATGCCCCATTTGATTCCTGCTAGTAC	250982
Qy	1800	AGGTGATTTAGCTTTTACGTTTCACTTTATATGGGTATACTCGAATPATAAATTTGGCGCTC	1859
Db	250983	AGGTGATTTAGCACTACGTTTCGACATTTTATGGTTATGATTTCTAATTTTATATGGAGATC	251042
Qy	1860	TATGTCATGGGACAAACGAAGTAGCATTTTAATAACGGATCAGGTTCTGGTGACGGTATCGA	1919
Db	251043	TATGTCATGGGACAAACGAAGTAGCATTTTAATAACGGATCAGGTTCTGGTGAGGTATCGA	251102

QY	1920	TAACCAAGTTGTTCTCTGAACAACTGATGAGCTGGTGAATTAACCAATTCAGAGGA	1979
Db	251103	TAACCAAGTTGTTCTCTGAACAACTGATGAGCTGGTGAATTAACCAATTCAGAGGA	251162
QY	1980	TTCCAGATTCTGACCCAGGTTTCAGATTCTGGCAGCGAATTTCTAAATTCAGATAGCGGTTTCAGA	2039
Db	251163	TTCCAGATTCTGACCCAGGTTTCAGATTCTGGCAGCGAATTTCTAAATTCAGATAGCGGTTTCAGA	251222
QY	2040	TTCCGGTAGTGATTTCTACATCAGATAGTGTTTCAGATTTCAGCGAGTGAATTCAGATTCAGC	2099
Db	251223	TTCTGGCAGTGATTTCTACATCAGATAGTGTTTCAGATTTCAGCGAGTGAATTCAGATTCAGC	251282
QY	2100	AAGTGATTTCAGACTCAGCGAGTGATTTCAGATTTCAGCA-----	2136
Db	251283	AAGTGATTTCAGACTCAGCGAGTGATTTCAGATTTCAGCAAGTGAATTCAGATTCAGCAAGTGA	251342
QY	2137	-----	2136
Db	251343	TTCCAGATTTCAGCAAGTGATTTCAGACTCAGCAAGTGATTTCAGATTTCAGCAAGTGAATTCAGA	251402
QY	2137	-----AGCGATTTCGACTTCAGCGAGCGAATTCGCACTCAGA	2171
Db	251403	TTCCAGCAAGCGAATTCAGATTTCAGCGAGCGAATTCAGATTTCAGCGAGCGAATTCAGATTTCAGC	251462
QY	2172	CAATGACTTCGGATTTCAGATAGCGAATTTCTGACTTCAGACAGTGAATTCAGATTTCGACAGTGA	2231
Db	251463	GAGTGATTTCGACTTCAGCGAGCGAATTCAGACTTCAGATAGTGACTTCAGATTTCGCACTCAGA	251522
QY	2232	CTCAGATTTCAGATAGCGAATTTCTGACTTCAGACAGTGAATTCAGATTTCAGATAGCGAATTCAGA	2291
Db	251523	TTCCGACTTCAGATAGCGAATTCAGATTTCAGACAGCGAATTTCTGACTTCAGACAGCGAATTCCTGA	251582
QY	2292	TTCCAGATAGCGAATTCAGATTTCGCAAGTGAATTCGCACTTCAGACAGCGAATTTCTGACTTCGCA	2351
Db	251583	CTCAGACAGTGAATTCAGATTTCGCAAGTGAATTCGCACTTCAGACAGTGAATTCAGATTTCGCA	251642
QY	2352	CAGTGATTTCGCACTTCAGAC-----AGCGATTTCAGATTTCGCAAGTGA	2393
Db	251643	TAGCGATTTCGCACTTCAGACAGTGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGA	251702
QY	2394	TTCCGACTTCAGATAGCGAATTCGCACTTCAGATAGCGAATTCAGATTTCAGACAGCGAATTCAGA	2453
Db	251703	TTCCGACTTCAGATAGCGAATTCGCACTTCAGATAGCGAATTCAGATTTCAGACAGCGAATTCAGA	251762
QY	2454	TTCCAGACAGCGAATTCAGATTTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCA	2513
Db	251763	TTCCAGACAGCGAATTCGACTTCAGACAGTGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCA	251822
QY	2514	CAGTGACTTCGGATTTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGA	2573
Db	251823	CAGTGATTTCAGACTTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGA	251882
QY	2574	CTCAGACTTCAGACAGTGAATTCGCAATTCAGCGAGTGAATTCGCAATTCAGATAGTGAATTCGCA	2633
Db	251883	TTCCAGATTTCGCAATTCAGATTTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGAATTCGCA	251942
QY	2634	TTCCGACAGTGAATTCGCAATTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCA	2693
Db	251943	TTCCGCGAGTGAATTCAGATTTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCA	252002
QY	2694	TAGCGAATTCGCAATTCAGATAGCGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGA	2753
Db	252003	CAGTGATTTCAGATTTCAGATAGTGAATTCAGATTTCGCAAGTGAATTCAGATTTCGCAAGTGA	252062
QY	2754	TTCCAGATTTCAGACAGC-----GACTTCAGACAGTGAATTCAGATTTCAGATAGTGAATTCGCA	2807
Db	252063	CTCCGATTTCAGACAGCGAATTCAGATTTCAGATAGTGAATTCAGATTTCGCAAGTGAATTCGCA	252122
QY	2808	TTCCAGCGAGTGAATTCAGACTTCAGTGAATTCGCAATTCAGATAGTGAATTCAGATAGTGAATTC	2858
Db	252123	TTCCAGCGAGTGAATTCAGACTTCAGTGAATTCGCAATTCAGATAGTGAATTCAGATAGTGAATTC	252182

QY 301 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGAAAAAATCGAATTCGGCTGGCTTCA 360
Db 1 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGAAAAAATCGAATTCGGCTGGCTTCA 60
QY 361 GTGCTTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAGAAGCAGATCAAGT 420
Db 61 GTGCTTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAGAAGCAGATCAAGT 120
QY 421 GAAAAATAGTTTACCGCAATCTGATAGCGCAAGTAAAGCAAGCAAGTAATGATTCAGT 480
Db 121 GAAAAATAGTTTACCGCAATCTGATAGCGCAAGTAAAGCAAGCAAGTAATGATTCAGT 180
QY 481 AGCGTTAGTGTGACCTTAAACAGACGACACAAACGTCAGTGAATCTAAAAACATCGTCA 540
Db 181 AGCGTTAGTGTGACCTTAAACAGACGACACAAACGTCAGTGAATCTAAAAACATCGTCA 240
QY 541 AACCTAATAATGGCGAAACGAGTGTGGGCHAAATCCAGCAACAGAGAAACGACAAA 600
Db 241 AACCTAATAATGGCGAAACGAGTGTGGGCHAAATCCAGCAACAGAGAAACGACAAA 300
QY 601 TCATCATCAAAATCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTAG 660
Db 301 TCATCATCAAAATCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTAG 360
QY 661 ACAAGCAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGCGAGAA 720
Db 361 ACAAGCAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGCGAGAA 420
QY 721 TTAGTGATCAACAAAGTAATGAACGACTTTAATGATCTAATACAGTATCATCTGTA 780
Db 421 TTAGTGATCAACAAAGTAATGAACGACTTTAATGATCTAATACAGTATCATCTGTA 480
QY 781 AATTCACTCAAAATCTTACAAATCGGAAATGTTTCAACAAACGCAAGTACTCAACT 840
Db 481 AATTCACTCAAAATCTTACAAATCGGAAATGTTTCAACAAACGCAAGTACTCAACT 540
QY 841 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 900
Db 541 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT 600
QY 901 GTAGTTAATCAAGCGTTAATACAGTGGCGCTAGATGAGAGCATTTAGTTTACGGGCA 960
Db 601 GTAGTTAATCAAGCGTTAATACAGTGGCGCTAGATGAGAGCATTTAGTTTACGGGCA 660
QY 961 GTAGCTCAGATGACACCGGAGCTGGCACAGATATTACGAATCAGTTGACCAATGTGACA 1020
Db 661 GTAGCTCAGATGACACCGGAGCTGGCACAGATATTACGAATCAGTTGACCAATGTGACA 720
QY 1021 GTTGGTATTCGACTGTGTTACGACTGTGTATCCGCAACCAAGCAGTTATGTCAAACTGAAT 1080
Db 721 GTTGGTATTCGACTGTGTTACGACTGTGTATCCGCAACCAAGCAGTTATGTCAAACTGAAT 780
QY 1081 TATGCGTTTTCAGTGCCTAATCTGCTGTTAAAGTGACACATTCAAATAACTGTACCT 1140
Db 781 TATGCGTTTTCAGTGCCTAATCTGCTGTTAAAGTGACACATTCAAATAACTGTACCT 840
QY 1141 AAAGAAATTAACCTTAAATGGTTAACTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA 1200
Db 841 AAAGAAATTAACCTTAAATGGTTAACTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA 900
QY 1201 GATCAAGTATTCGCAATGTGTAAATCGATAGTGATGGTAATGTTATTTATACATTTACA 1260
Db 901 GATCAAGTATTCGCAATGTGTAAATCGATAGTGATGGTAATGTTATTTATACATTTACA 960
QY 1261 GACTATCTAAATACTAAAGATGATTAAGACCACTTTGACCATCGCCGCTTATTTGAC 1320
Db 961 GACTATCTAAATACTAAAGATGATTAAGACCACTTTGACCATCGCCGCTTATTTGAC 1020
QY 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGGCATAGGTAGTACAACA 1380
Db 1021 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGCTACTGGCATAGGTAGTACAACA 1080

QY 1381 GCAACCAAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTATTAACCTATCTATT 1440
Db 1081 GCAACCAAAACAGTATTAGTAGATTATGAAAAATATGGTAAGTTTATTAACCTATCTATT 1140
QY 1441 AAAGGTACAATTGACCAAAATCGATAAACAACAAATAATACGTATCGTCAGACAAATTTATGTC 1500
Db 1141 AAAGGTACAATTGACCAAAATCGATAAACAACAAATAATACGTATCGTCAGACAAATTTATGTC 1200
QY 1501 AATCCAAGTGGAGATAAGCTTTATTCGCGGGGTTTTAAACAGGTAATTTAAAAACAAATAG 1560
Db 1201 AATCCAAGTGGAGATAAGCTTTATTCGCGGGGTTTTAAACAGGTAATTTAAAAACAAATAG 1260
QY 1561 GATAGTAATGATTAATAGATCAGCAAAATACAAGTATTAAGTATATAAAGTAGATAAT 1620
Db 1261 GATAGTAATGATTAATAGATCAGCAAAATACAAGTATTAAGTATATAAAGTAGATAAT 1320
QY 1621 GAGCTGATTTATCTGAAAAGTTTACTTTTGTGAATCCAGAAAACTTTGAGGATCTCACTAAT 1680
Db 1321 GAGCTGATTTATCTGAAAAGTTTACTTTTGTGAATCCAGAAAACTTTGAGGATCTCACTAAT 1380
QY 1681 AGTGTGAATATTACANTCCCAATCCCAATCAATATAAAGTAGAGTTTAAATACGCTGAT 1740
Db 1381 AGTGTGAATATTACANTCCCAATCCCAATCAATATAAAGTAGAGTTTAAATACGCTGAT 1440
QY 1741 GATCAAAATTACAAACACCGTATATAGTAGTTTGTAAATGCTCATATTTGATCCGAAATAGCAAA 1800
Db 1441 GATCAAAATTACAAACACCGTATATAGTAGTTTGTAAATGCTCATATTTGATCCGAAATAGCAAA 1500
QY 1801 GGTGATTTAGCTTTTACGTTTCAACTTTATATCGGTATTAATCGAAATATAAATTTGGCGCTCT 1860
Db 1501 GGTGATTTAGCTTTTACGTTTCAACTTTATATCGGTATTAATCGAAATATAAATTTGGCGCTCT 1560
QY 1861 ATGTCATGGGACACAGTAGTAGCTTTAATAACCGATCAGGTTCTGTTGAGCGGATTCGAT 1920
Db 1561 ATGTCATGGGACACAGTAGTAGCTTTAATAACCGATCAGGTTCTGTTGAGCGGATTCGAT 1620
QY 1921 AAACAGTTGTTCTGAAACAACTGATGAGCTCGTGAAATTTGAACCAATTTCCAGAGAT 1980
Db 1621 AAACAGTTGTTCTGAAACAACTGATGAGCTCGTGAAATTTGAACCAATTTCCAGAGAT 1680
QY 1981 TCAGATTCGACCCAGGTTTCAGATTTCTGGCAGCGATTCTAATTCAGATAGCGTTTCAGAT 2040
Db 1681 TCAGATTCGACCCAGGTTTCAGATTTCTGGCAGCGATTCTAATTCAGATAGCGTTTCAGAT 1740
QY 2041 TCGGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGTATTCAGATTCAGCA 2100
Db 1741 TCGGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGTATTCAGATTCAGCA 1800
QY 2101 AGTGATTCAGACTCAGCGAGTGTATTCAGATTCAGCAAGCGATTCCGACTCAGCGAGCGAT 2160
Db 1801 AGTGATTCAGACTCAGCGAGTGTATTCAGATTCAGCAAGCGATTCCGACTCAGCGAGCGAT 1860
QY 2161 TCCGACTCAGACAAATGATCGGATTCAGATAGCGATTCTGACTCAGACAGTGTATTCAGAT 2220
Db 1861 TCCGACTCAGACAAATGATCGGATTCAGATAGCGATTCTGACTCAGACAGTGTATTCAGAT 1920
QY 2221 TCCGACTCAGACAAATGATCGGATTCAGATAGCGATTCTGACTCAGACAGTGTATTCAGAT 2280
Db 1921 TCCGACTCAGACAAATGATCGGATTCAGATAGCGATTCTGACTCAGACAGTGTATTCAGAT 1980
QY 2281 AGCGATTCAGATTCAGATAGCGATTTCAGATTCGCAAGCGATTCCGACTCAGACAGCGAT 2340
Db 1981 AGCGATTCAGATTCAGATAGCGATTTCAGATTCGCAAGCGATTCCGACTCAGACAGCGAT 2040
QY 2341 TCTGACTCCGACAGTANTCCGACTCAGACAGCGATTTCAGATTCGCAAGCGATTCCGACT 2400
Db 2041 TCTGACTCCGACAGTANTCCGACTCAGACAGCGATTTCAGATTCGCAAGCGATTTCGACT 2100
QY 2401 TCAGATAGCGATTCGCGACTCAGATAGCGATTTCAGATTCGCAAGCGATTTCAGATTCAGAC 2460
Db 2101 TCAGATAGCGATTCGCGACTCAGATAGCGATTTCAGATTCGCAAGCGATTTCAGATTCAGAC 2160
QY 2461 AGCGATTCAGATTCAGATAGCGATTTCAGATTCGCAAGCGATTTCAGATTCGCAAGCGAT 2520

|||||
2161 AGCGATTTCAGATTTCAGATAGCGATTTCAGATTCGGACAGTGCAT 2220
|||||
2521 TCGGATTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCAT 2580
|||||
2221 TCGGATTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCAT 2280
|||||
2581 TCAGACAGTGCATTTCGGATTTCAGATAGCGATTTCGGACAGTGCATTCGGACAGTGCAT 2640
|||||
2281 TCAGACAGTGCATTTCGGATTTCAGATAGCGATTTCGGACAGTGCATTCGGACAGTGCAT 2340
|||||
2641 AGTGCATTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2700
|||||
2341 AGTGCATTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2400
|||||
2701 TCGGACTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2760
|||||
2401 TCGGACTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2442
|||||
2761 TCAGACAGTGCATTTCGGATTTCAGATAGCGATTTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2820
|||||
2443 TCAGACAGTGCATTTCGGATTTCAGATAGCGATTTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2502
|||||
2821 TCAGACTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2880
|||||
2503 TCAGACTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2562
|||||
2881 AGCGATTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2940
|||||
2563 AGCGATTTCAGATAGCGATTTCAGATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCATTCGGACAGTGCAT 2622
|||||
2941 ACTAATGCTTCTAATAAATAGGCTAAGATAGTAAAGAACATTCACAGATACAGT 3000
|||||
2623 ACTAATGCTTCTAATAAATAGGCTAAGATAGTAAAGAACATTCACAGATACAGT 2682
|||||
3001 TCTGAAGATGAAGCAAAATAGCTCACTAAATTTGGGGATTATTAGCATCAATPAGTTTCATTA 3060
|||||
2683 TCTGAAGATGAAGCAAAATAGCTCACTAAATTTGGGGATTATTAGCATCAATPAGTTTCATTA 2742
|||||
3061 CTACTTTTCAGAGAGAAAAAAGAAAAATAAAGATAAGAAA 3099
|||||
2743 CTACTTTTCAGAGAGAAAAAAGAAAAATAAAGATAAGAAA 2781
|||||

RESULT 12

BX571856_08

WPCOMMENT

Sequence split into 29 fragments LOCUS BX571856 Accession BX571856

Fragment Name	Begin	End
BX571856_00	1	110000
BX571856_01	100001	210000
BX571856_02	200001	310000
BX571856_03	300001	410000
BX571856_04	400001	510000
BX571856_05	500001	610000
BX571856_06	600001	710000
BX571856_07	700001	810000
BX571856_08	800001	910000
BX571856_09	900001	1010000
BX571856_10	1000001	1110000
BX571856_11	1100001	1210000
BX571856_12	1200001	1310000
BX571856_13	1300001	1410000
BX571856_14	1400001	1510000
BX571856_15	1500001	1610000
BX571856_16	1600001	1710000
BX571856_17	1700001	1810000
BX571856_18	1800001	1910000
BX571856_19	1900001	2010000
BX571856_20	2000001	2110000
BX571856_21	2100001	2210000
BX571856_22	2200001	2310000
BX571856_23	2300001	2410000

BX571856_24	2400001	2510000
BX571856_25	2500001	2610000
BX571856_26	2600001	2710000
BX571856_27	2700001	2810000
BX571856_28	2800001	2902619
Continuation (9 of 29) of BX571856 from base 800001 (BX571856 Staphylococcus aureus subst		
Query Match 76.8%; Score 2686.8; DB 1; Length 110000;		
Best Local Similarity 84.9%; Pred. No. 0;		
Matches 3216; Conservative 0; Mismatches 282; Indels 290; Gaps 7;		
QY	1	GGTACCAATAATTCACATCTGCTTTTGGAAAAAATATGATTTCAAGCTAGGATTACATTA 60
DB	87937	GGTACCAATAATTCACATCTGCTTTTGGAAAAAATATGATTTCAAGCTAGGATTACATTA 87996
QY	61	GGTAGAGTTTCATTAATAATAAATAAATGTTTGCATCAATCAATCGTACGTTGCTGTTTGA 120
DB	87997	GGCAGAGTTTCATTAATAATAAATAAATGTTTGCATCAATCAATCGTACGTTGCTGTTTGA 88056
QY	121	ATTCTTAAAAATAG-CAATAAAATAAAATGTTTGTAGTAAAGTATTATTGTGGATAATAA 179
DB	88057	ATTCTTAAAAACAGAGAAATGAAATGAAATGTTTGTAGTAAATATTATTGTAGTTAATAA 88116
QY	180	ATATCGATACAAATTAATTGCTTATTAATGCAATTTTGTGTATTAATTCATTAACAGAT 239
DB	88117	ATTCTAAAAACAAATTAATTGCTTATTAATGCAATTTTGTGTATTAATTCATTAACAGAT 88176
QY	240	TAAATATATCTT-AAAGGTTATATAGTTTAAATATAAATGACTTTTAAAAAGAGGGAATA 298
DB	88177	TAAATATATATTAATAAAGGTTATACAGTTTAAATATAAATGACTTTTAAAAAGAGGGAATA 88236
QY	299	AAATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTTGGCGTGGCTT 358
DB	88237	AAATGAATATGAAGAAAAACAGCAATTCGAAAAAATCGATTTGGCGTGGCTT 88296
QY	359	CAGTGTCTTAGTGATCGTTTAAATCGGTTTGGACTACTCAGCAGTAAAGACAGATGCAA 418
DB	88297	CAGTGTCTTAGTGATCGTTTAAATCGGTTTGGACTACTCAGCAGTAAAGACAGATGCAA 88356
QY	419	GTGAAAATAGTTTACGCAATCTGATAGCGCAAGTAAACGAAACGAAAGTAATGATTCAA 478
DB	88357	GTGAAAATAGTATGACACAAACGGAATAACGAGTAAATGAGACAAAGTAATGATTCCAA 88416
QY	479	GTAGCGTTAGTGTGCACCTTAAAAACAGACACAAACGTCGATGATCTATAAACATCGT 538
DB	88417	GTAGCGTTAATGCTGCACCTTAAAAACAGACACAAACGTCGATGATCTATTAACAACGA 88476
QY	539	CAAACTAATAATGCGCAAAACGAGTGTGCGCAAAATCCAGCAACACAGCAACGACAC 598
DB	88477	CAAACTAATAATGCGCAAAACGAGTGTGCGCAAAATCCAGCAACACAGCAACGACAC 88536
QY	599	AATCATCATCAAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTA 658
DB	88537	AATCAGCATCAAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAGTACTACTA 88596
QY	659	CGACAAACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATTCGGAGG 718
DB	88597	CGCNAACGAATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATTCGGAGG 88656
QY	719	AATTAGTGAATCAAAACGAATGAACGACTTTTATGATCTAATACAGTATCATCTG 778
DB	88657	AATCAGTGAATCAAAACGAATGAACGACTTTTATGATCTAATACAGTATCATCTG 88716
QY	779	TAAATTCACCTCAAAATCTTACAAATCGGAAATGTTTCAACAAACGAAAGTACTTCAA 838
DB	88717	TAAATTCACCTCAAAATCTTACAAATCGGAAATGTTTCAACAAACGAAAGTACTTCAA 88776
QY	839	CTGAAGCAACACCTTCAAAACGAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAG 898
DB	88777	CTGAAGCAACACCTTCAAAACGAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAG 88836
QY	899	ATGTAGTTAATCAAGCGGTTTAAATACAGTGGCGCTAGATGAGAGCATTTAGTTTAGCGG 958

88937	ATGTAGTTAATCAAGCGGTTAAATACAAGTGCCTGAGATGAGAGCATTATAGTTTTAGCGG	88899
959	CAGTAGCTGCAGATGACCGCGCAGCTGGCACAGATATATACGAATCAGTTTGACGAATGTGA	1018
88897	CTGTAGCTGCAGATGACCGCGCTGCTGGCAAGAGATATATAGAAATCAGTTTGACGAATGTGA	88956
1019	CAGTTGGTATGACTCTGTGTAACGACTGTGTATCCGCACCAAGCAGGTTATGTCAAACTGA	1078
88957	CAGTTGGTATGACTCTGTGAGATACACAGTTTATCCGCACCAAGCAGGCTATGTCAAACTGA	89016
1079	ATTATGCTGTTTTCAGTGCCTTAATCTGCTGTTTAAAGGTGACACATTCAAAATACTGTAC	1138
89017	ATTATGGGTTCTCAGTACCNAATGAGCGCTGTTCAAGGTGACACATTCAAAATACTGTGC	89076
1139	CTAAAGAAATTAACCTTAATGGTGAACCTTCAACTCTAAAGTGCCACCAATATGCGCTG	1198
89077	CCAAAGAAATTAACCTTAATGGTGAACCTTCAACTCTAAAGTGCCACCAATATGCGCGG	89136
1199	GAGATCAAGTATTTGGCAAAATGGTGTAATCGATAGTGATGTAATGTTATTTATACATTTA	1258
89137	GAGATCAAGTATTTGGCAAAATGGTGTAATCGATAGTGATGTAATGTTATTTATACATTTA	89196
1259	CAGACTATGTAANATCTAAGATGATGTAAGAGCAACTTTGACCAATGCCCGTTATATGT	1318
89197	CAGACTATGTAANATCTAAGATGATGTAAGAGCAACTTTGACCAATGCCCGTTATATGT	89256
1319	ACCTGAAATGTTAAAAAGACAGGTAATGTGCACATTCGCTACTCGCATAGGTAGTACAA	1378
89257	ACCTGAAATGTTAAAAAGACAGGTAATGTGCACATTCGCTACTCGCATAGGTAGTACAA	89316
1379	CAGCAAAACAAAACAGTATTTAGTAGATTTATGAAAAATATGGTAAGTTTATTAACCTTATCTA	1438
89317	CAGCAAAACAAAACAGTATTTAGTAGATTTATGAAAAATATGGTAAGTTTATTAACCTTATCTA	89376
1439	TTAAAGGTCAAAATTGACCAAAATCGATAAAACAAATAATACGTAACGTAAGTTTATG	1498
89377	TTAAAGGTCAAAATTGACCAAAATCGATAAAACAAATAATACGTAACGTAAGTTTATG	89436
1499	TCAATCCAAGTGAGATAACGTTTATTCGCGCGGTTTTAACAGGTAATTTTAAACCAAAATA	1558
89437	TCAATCCAAGTGAGATAACGTTTATTCGCGCGGTTTTAACAGGTAATTTTAAACCAAAATA	89496
1559	CGGATAGTAATGCATTTAATPAGATCAGCAAAAATACAAGTATTTAAAAGTATATAAAGTAGATA	1618
89497	CGGATAGTAATGCATTTAATPAGATCAGCAAAAATACAAGTATTTAAAAGTATATAAAGTAGATA	89556
1619	ATCCAGCTCAATTAICTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGAGATGCTACTA	1678
89557	ATCCAGCTCAATTTGCTGAAAGTTTATTATGTGAATCCAGATAAACTTTGAAGATGCTACTG	89616
1679	ATAGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGTAGGTTTAAATCGCGCTG	1738
89617	ATAGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGTAGGTTTAAATCGCGCTG	89676
1739	ATGATCAAAATTACACACCGTATATPAGTAGTTGTTAATGGTCAATTTGATCCGAATAGCA	1798
89677	ATGATCAAAATTACACACCATATATPAGTTGTTAATGGTCAATTTGATCCGAATAGTA	89736
1799	AGGTGATTTTACCTTTACCTTCACTTTATATGGGTATAACTCGAATATAATTTGGCGCT	1858
89737	AGGTGATTTTACCTTTTACCTTCACTTTATATGGGTATAAATTCGAAATAAATTTGGCGAT	89796
1859	CTATGTCTATGGACAAACGAGTATGATTTAATACGGATCAGGTTCTGTTGAGCGGTATCG	1918
89797	CAATGTCTATGGATTAATGAGTAGCAATTTAATAACGGATCAGGTTCTGTTGAGCGGTATCG	89856
1919	ATAAACCGATGTTCTCTGAACAACTGATGAGCCCTGGTGAATTTGAACCAATTTCCAGAGG	1978
89857	ATAAACCTGTTCTCTGAACAACTGATGAGCCGGGTGAATTTGAACCAATTTCCAGAGG	89916
1979	ATTTCAGATTTCTACCCA-----GGTTCAGATTTCTGGCAGCGATTTAAATTCAG	2026
89917	ATTTCAGATTTCTACCCAGGTTTCAGATAGTGGTTTCAGATTTCTGGCAGCGATTTAAATTCAG	89976

```
QY 2833 --AGTGACTCGGATTCATCAAGTGATTCGACTCAGAAAGTGATTCAAATAGCGATTCG 2890
Db 91057 CGAGTGACACAGGATCAGATACGACTCAGACTCAGAAAGTGATTCGATAGCGATTCG 91116
QY 2891 AGTCAGGTTCTAACAAATAGTGTTCGCGCTTAATTCACCTAAATGGTACTAATGCTT 2950
Db 91117 ATTCAAGTTCTAACAAATAGTGTTCGCGCTTAATTCACCTAAATGGTACTAATGCTT 91176
QY 2951 CTAATAAAATGAGGCTAAAGATAGTAAGAACCAATACCAGATACAGGTTCTGAAGATG 3010
Db 91177 CAAATAAAATGAGGCTAAAGATAGTAAGAACCAATACCAGATACAGGTTCTGAAGATG 91236
QY 3011 AAGCAAAATACGTCACCTAATTTGGGATTAATAGCATCAATAGGTTCAATACCTACCTTTCA 3070
Db 91237 AAGCTAATACGTCACCTAATTTGGGATTAATAGCATCAATAGGTTCAATACCTACCTTTCA 91296
QY 3071 GAAGAAAAAGAAAAATAGAAATAGAAATAGTAATAGTAATAGTAATAGTAATAGTAATAG 3130
Db 91297 GAAGAAAAAGAAAAATAGAAATAGAAATAGTAATAGTAATAGTAATAGTAATAGTAATAG 91356
QY 3131 ATTCAATGAAGAACCACTTAAGAGTGCTCTTTTACTTCTGGATTTTCCAAATATATGCT 3190
Db 91357 ATTCATGAAGAACCACTTAAGAGTGCTCTTTTACTTCTGGATTTTCCAAATATATGCT 91416
QY 3191 TTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3250
Db 91417 TTGAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 91476
QY 3251 ATAATTTGGCTTGGCGAAAAATAGGTTGAAGTGAAGTGTGTTAATTTAGGGAATAATAG 3310
Db 91477 ATAATTTGGCTTGGCGAAAAATAGGTTGAAGTGAAGTGTGTTAATTTAGGGAATAATAG 91536
QY 3311 GAGAAAAATACAGTTGAAAAATTAATGCTAGTATTTATCATTTGGGAGCATTTGCTATCA 3370
Db 91537 GAGAAAAATACAGTTGAAAAATTAATGCTAGTATTTATCATTTGGGAGCATTTGCTATCA 91596
QY 3371 CAAATTTGGGAAAGTAATCGTCGAGTGCAGTGGTTCCTGGGAGAAAGTAATCCATATGTA 3430
Db 91597 CAAATTTGGGAAAGTAATCGTCGAGTGCAGTGGTTCCTGGGAGAAAGTAATCCATATGTA 91656
QY 3431 TCTGAGTCTGTTGAACTGACTAATAATAAAAAATAATCTAGAACAGTGAAGAGTATAAG 3490
Db 91657 TCTGAGTCTGTTGAACTGACTAATAATAAAAAATAATCTAGAACAGTGAAGAGTATAAG 91716
QY 3491 AAAAGCTT 3498
Db 91717 GAAAAATTT 91724

RESULT 13
LOCUS AX343875 2792 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 3 from Patent WO0198499.
ACCESSION AX343875
VERSION AX343875.1 GI:18491941
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1 Staphylococcus aureus
AUTHORS Bacteria; Firmicutes; Bacillales; Staphylococcus.
TITLE Foster S.D., McDowell, P.D., Brummell, K.D. and Clarke, S.D.
JOURNAL Antigenic polypeptides
PATENT Patent: WO 0198499-A 3 27-DEC-2001;
JOURNAL THE UNIVERSITY OF SHEFFIELD (GB); Biosynexus Incorporated (US)
FEATURES
    Location/Qualifiers
    1..2792
        /organism="Staphylococcus aureus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:11280"

ORIGIN
Query Match 50.8%; Score 1776.4; DB 6; Length 2792;
Best Local Similarity 99.9%; Pred. No. 1.7e-257;
```

```
Matches 1788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GGTACCAATAAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTAACATTA 60
Db 1003 GGTACCAATAAATACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGGATTAACATTA 1062
QY 61 GGTAGAGTTTCATATTAATAAATAAATAATGTTTGCAATCAAAATCGTACGTTGCTGTTGTA 120
Db 1063 GGTAGAGTTTCATATTAATAAATAAATAATGTTTGCAATCAAAATCGTACGTTGCTGTTGTA 1122
QY 121 ATTCTTAAATAGCAATTAATAAATAAATAATGTTTGTTAGTAAGTATTAATTTGGTGGATTAATAA 180
Db 1123 ATTCTTAAATAGCAATTAATAAATAAATAATGTTTGTTAGTAAGTATTAATTTGGTGGATTAATAA 1182
QY 181 TATCGATACAAATTAATTTGCTATTAATGCAATTTTATGTTATTAATTTCCATTAACAGAGATT 240
Db 1183 TATCGATACAAATTAATTTGCTATTAATGCAATTTTATGTTATTAATTTCCATTAACAGAGATT 1242
QY 241 AAATATATC-TTAAAGGGTATATAGTTTAATATAAATCACTTTTAAAAAGAGGGAATAA 299
Db 1243 AAATATATC-TTAAAGGGTATATAGTTTAATATAAATCACTTTTAAAAAGAGGGAATAA 1302
QY 300 AATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTGGCTTC 359
Db 1303 AATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTGGCTTC 1362
QY 360 AGTGCTTGTAGTACGTTAAATCGGTTTGGGACTACTCAGCAGTAAGAAAGCAGATGCAAG 419
Db 1363 AGTGCTTGTAGTACGTTAAATCGGTTTGGGACTACTCAGCAGTAAGAAAGCAGATGCAAG 1422
QY 420 TGAATAATAGTGTACGCAATCTGATPAGCGCAAGTAACGAAAAAGAAAGTAATGATTCAAG 479
Db 1423 TGAATAATAGTGTACGCAATCTGATPAGCGCAAGTAACGAAAAAGAAAGTAATGATTCAAG 1482
QY 480 TAGCGTTAGTGTGCAACCTTAATAACAGACACACAAACGTGAGTGATCTAATAACATCGTC 539
Db 1483 TAGCGTTAGTGTGCAACCTTAATAACAGACACACAAACGTGAGTGATCTAATAACATCGTC 1542
QY 540 AAACACTAATTAATGCGAAAAACGAGTGTGCGCAAAATCCAGCACAAACAGGAAACGACACA 599
Db 1543 AAACACTAATTAATGCGAAAAACGAGTGTGCGCAAAATCCAGCACAAACAGGAAACGACACA 1602
QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 659
Db 1603 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 1662
QY 660 GACAAAGAAATCAAGCTAATATACACCGGCAACACTCAATCAAGCAATCAAAATGCGGAGCA 719
Db 1663 GACAAAGAAATCAAGCTAATATACACCGGCAACACTCAATCAAGCAATCAAAATGCGGAGCA 1722
QY 720 ATTAGTGAATCAACCAAGTAATGAAACGACTTTTAAATGATCTAATACAGTATCACTCTGT 779
Db 1723 ATTAGTGAATCAACCAAGTAATGAAACGACTTTTAAATGATCTAATACAGTATCACTCTGT 1782
QY 780 AAATTCACCTCAAAATTTCTACAAATGCGGAAAAATGTTTCAACAAACGCAAGATCTTCAAC 839
Db 1783 AAATTCACCTCAAAATTTCTACAAATGCGGAAAAATGTTTCAACAAACGCAAGATCTTCAAC 1842
QY 840 TGAAGCAACACCTTCAACCAATGAATCAGCTCCACAGAGTACAGATGCAAGTGAAGTAAGA 899
Db 1843 TGAAGCAACACCTTCAACCAATGAATCAGCTCCACAGAGTACAGATGCAAGTGAAGTAAGA 1902
QY 900 TGTAGTGAATCAAGCGGTTAATACAGTGGCTAGAAATGAGAGCAATTTAGTTTACGGCG 959
Db 1903 TGTAGTGAATCAAGCGGTTAATACAGTGGCTAGAAATGAGAGCAATTTAGTTTACGGCG 1962
QY 960 AGTAGTGCAGATGCAACCGGAGCTGGCAGATATTAAGAAATCAGTTGACGAATGTGAC 1019
Db 1963 AGTAGTGCAGATGCAACCGGAGCTGGCAGATATTAAGAAATCAGTTGACGAATGTGAC 2022
QY 1020 AGTTGGTATTGACTCTGGTACGACTGTGTATCGGACCAAGCAGGTTATGTCAAACCTGAA 1079
Db 2023 AGTTGGTATTGACTCTGGTACGACTGTGTATCGGACCAAGCAGGTTATGTCAAACCTGAA 2082
```

```
Qy 1080 TTATGGTTTTTCAGTGCTAATCTCTGCTGTTAAAGGTGACACATTCAAAATAAAGTGTACC 1139
Db 2083 TTATGGTTTTTCAGTGCTAATCTCTGCTGTTAAAGGTGACACATTCAAAATAAAGTGTACC 2142
Qy 1140 TAAAGAAATTAACCTTAATAGGTGTAACCTCACTGCTAAAGTGCCACCAATTAATGGCTGG 1199
Db 2143 TAAAGAAATTAACCTTAATAGGTGTAACCTCACTGCTAAAGTGCCACCAATTAATGGCTGG 2202
Qy 1200 AGATCAAGTATGTCGAATGCTGTAATCGATAGTGGTGAATGTTATTTATATCAATTTAC 1259
Db 2203 AGATCAAGTATGTCGAATGCTGTAATCGATAGTGGTGAATGTTATTTATATCAATTTAC 2262
Qy 1260 AGACTATGTAATTAACCTTAATAGGTGTAACCTCACTGCTAAAGTGCCACCAATTAATGGCTGG 1319
Db 2263 AGACTATGTAATTAACCTTAATAGGTGTAACCTCACTGCTAAAGTGCCACCAATTAATGGCTGG 2322
Qy 1320 CCCTGAAATATGTTAAAGACAGTAATGTGACATTCGGCTACTGGCATAGTAGTACAAC 1379
Db 2323 CCCTGAAATATGTTAAAGACAGTAATGTGACATTCGGCTACTGGCATAGTAGTACAAC 2382
Qy 1380 AGCAAAACAAACAGTATTAGTAGATTTATGAAAATATGGAATATGGAATATGGAATATGGAAT 1439
Db 2383 AGCAAAACAAACAGTATTAGTAGATTTATGAAAATATGGAATATGGAATATGGAATATGGAAT 2442
Qy 1440 TAAAGGTACAAATTCGACAAATTCGATAAAACAAATTAATACGTTATCGTCAGACAAATTTATGT 1499
Db 2443 TAAAGGTACAAATTCGACAAATTCGATAAAACAAATTAATACGTTATCGTCAGACAAATTTATGT 2502
Qy 1500 CAATCCAAGTGGAGATAACGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAAATAC 1559
Db 2503 CAATCCAAGTGGAGATAACGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAAATAC 2562
Qy 1560 GGATAGTAATGCATTAATAGATCAGCAAAATACAAATTAATTAAGTATTAAGTAGATAA 1619
Db 2563 GGATAGTAATGCATTAATAGATCAGCAAAATACAAATTAATTAAGTATTAAGTAGATAA 2622
Qy 1620 TGCAAGTGAATTTACTGAAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCATAA 1679
Db 2623 TGCAAGTGAATTTACTGAAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCATAA 2682
Qy 1680 TAGTGCAATTTACATTCGAAATCCAAATCAATTAATTAAGTAGAGTTAATACGCTGA 1739
Db 2683 TAGTGCAATTTACATTCGAAATCCAAATCAATTAATTAAGTAGAGTTAATACGCTGA 2742
Qy 1740 TGATCAAAATTAACACCGGTATATAGTAGTTGTTAATGGTCATATGATC 1789
Db 2743 TGATCAAAATTAACACCGGTATATAGTAGTTGTTAATGGTCATATGATC 2792
```

```
RESULT 14
AX768506 2792 bp DNA linear PAT 02-JUL-2003
LOCUS
DEFINITION Sequence 92 from Patent WO03011899.
ACCESSION AX768506
VERSION AX768506.1 GI:32437047
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Foster, S., Mond, J., Clarke, S., McDowell, P. and Brummel, K.
TITLE Antigenic polypeptides
JOURNAL Patent: WO 03011899-A 92 13-FEB-2003;
THE UNIVERSITY OF SHEFFIELD (GB); Biosynex Incorporated (US)
FEATURES
source
1. .2792
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:1280"
```

Query Match 50.8%; Score 1776.4; DB 6; Length 2792;

Best Local Similarity 99.9%; Pred. No. 1.7e-257;
Matches 1788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy 1 GGTACCAATAATTCACATCTGCTTTTGAATAATATGATTTCAAGTAGATTAACATTA 60
Db 1003 GGTACCAATAATTCACATCTGCTTTTGAATAATATGATTTCAAGTAGATTAACATTA 1062
Qy 61 GGTAGAGTTCATATTAATAATAAATAATGTTTCAATCAATCGTACGTTGTCGTTGTA 120
Db 1063 GGTAGAGTTCATATTAATAATAAATAATGTTTCAATCAATCGTACGTTGTCGTTGTA 1122
Qy 121 ATTCTTTAAATAGCAATTAATAATAAATAATGTTTGTAGTAAAGTATTTATTTGTGATATAAAA 180
Db 1123 ATTCTTTAAATAGCAATTAATAATAAATAATGTTTGTAGTAAAGTATTTATTTGTGATATAAAA 1182
Qy 181 TATCGATACAAATTAATGCTATTAATGCAATTTTGTGTAATAATCCATTAACAGAGATT 240
Db 1183 TATCGATACAAATTAATGCTATTAATGCAATTTTGTGTAATAATCCATTAACAGAGATT 1242
Qy 241 AATATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTAAATAAGAGGAATAA 299
Db 1243 AATATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTAAATAAGAGGAATAA 1302
Qy 300 AATGAATATCAAGAAAAAGAAAAACACGCAATTCGGAAAAAATCGAATTTGGCGTGGCTTC 359
Db 1303 AATGAATATCAAGAAAAAGAAAAACACGCAATTCGGAAAAAATCGAATTTGGCGTGGCTTC 1362
Qy 360 AGTGCCTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAGAAGCAGATGCAAG 419
Db 1363 AGTGCCTGTAGGTACGTTAATCGGTTTGTGACTACTCAGCAGTAAGAAGCAGATGCAAG 1422
Qy 420 TGAATAATAGTGTACGCAATCTGTATAGCGCAAGTAACGAAAGCAAAAGTAATGATTTCAAG 479
Db 1423 TGAATAATAGTGTACGCAATCTGTATAGCGCAAGTAACGAAAGCAAAAGTAATGATTTCAAG 1482
Qy 480 TAGCGTTAGTGTGCACTTAAACACAGACGACAAACGTCAGTGATGATACTAAAAATCATCGTC 539
Db 1483 TAGCGTTAGTGTGCACTTAAACACAGACGACAAACGTCAGTGATGATACTAAAAATCATCGTC 1542
Qy 540 AATACACTAATAATGGCGAAACGAGTGTGGCGAAATCCAGACACACAGAAACGACACA 599
Db 1543 AATACACTAATAATGGCGAAACGAGTGTGGCGAAATCCAGACACACAGAAACGACACA 1602
Qy 600 ATCATCATCAACAAATTCGAATACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 659
Db 1603 ATCATCATCAACAAATTCGAATACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 1662
Qy 660 GACAAAGAAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATAAATCGGAGGA 719
Db 1663 GACAAAGAAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATAAATCGGAGGA 1722
Qy 720 ATTAGTGAATCAACAAAGTAATGAAACGACTTTTAAATGATTAATACAGTATCATCTGT 779
Db 1723 ATTAGTGAATCAACAAAGTAATGAAACGACTTTTAAATGATTAATACAGTATCATCTGT 1782
Qy 780 AAATTCACCTCAAAATTTCTAATAATGGGAAATGTTTCAACCAACGCAAGATACTTCAAC 839
Db 1783 AAATTCACCTCAAAATTTCTAATAATGGGAAATGTTTCAACCAACGCAAGATACTTCAAC 1842
Qy 840 TGAAGCAACACCTTCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 899
Db 1843 TGAAGCAACACCTTCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1902
Qy 900 TGTAGTGAATCAAGCGGTTAATAACAGTGGCTGCTAGAAATGAGAGCAATTTAGTTAGCGGC 959
Db 1903 TGTAGTGAATCAAGCGGTTAATAACAGTGGCTGCTAGAAATGAGAGCAATTTAGTTAGCGGC 1962
Qy 960 AGTAGCTGAGATGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
Db 1963 AGTAGCTGAGATGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2022
Qy 1020 AGTTGGTATGACTCTGGTACGACTGTGTATCGGACCAAGCAGGTTATGTCAAACTGAA 1079
```


Db 658 TAGCAAAAGTTGAAGCAACAAATACAGATAAATAAGCTTACTCAGACTTCAGACATATCGA 717
Qy 1001 ATCAGTTGACGAATGTCAGAGTTGGTATTCGACTCTGGTACGACTGTGTATCCGACCAAG 1060
Db 718 ATAAACTCTCAAAATGTAACAGCAGCAATTTGAAGCGGAGATACGATTTACCCACATAAAG 777
Qy 1061 CAGGTTATGTCAAACTCAAAATTAATGTTTTCAGTGCCTAATCTCTGCTGTTTAAAGGTGACA 1120
Db 778 CAGATATGTAATTTAAATATTCGTTTCAAGCCCGCAGATGATTTCAAGCAGGAGATA 837
Qy 1121 CATTCAAAATPACTGTACCTPAAGAAATTAACCTTAAATGTTGTAACCTTCAACTGCTAAAG 1180
Db 838 GTATTTAAATTTACTATTCCTCAAGCATTAAACTTAAATGTTGTAACAGCTACTCTCTAAAG 897
Qy 1181 TGCACCAATATGCTGGAGATCAAGTATTTGGCAATGTTGTAATCGATAGTATGGTA 1240
Db 898 CAGCTATATTTATGGCAGGTGACCAATATTTAGCTACTGGGCAATCGATGAAGAGGAA 957
Qy 1241 ATGTTATTTATACATTTACAGACTATGTAATATCTAAAGATGATGTAAGGCAACTTTGA 1300
Db 958 ATTAAATTTATACGTTTACAGATTATGTAACATAAATAATATTTACAGGACAGATAT 1017
Qy 1301 CCATGCCCGCTTATTTGACCCCTGAAATGTTTAAAGACAGGTAAATGTCACATTTGGCTA 1360
Db 1018 CAATTCCTGGGTATATTGATCTCTAAATAATGTAACACATACGGGTAAAGTCAATTTAGAGA 1077
Qy 1361 CTGGCATAGGTAGTACAAAGCAGCAAAACAGCTATTAGTAGATTATGAAATATGTTA 1420
Db 1078 CATCTATTTGGCCAAACAACTGCAAGGAAACCTGTGACAGTTGATTATGAAAGTATGGAG 1137
Qy 1421 AGTTTATTAACCTTATCTATTAAAGGTACAATTGACCAAAATCGATAAAACAAATAATACGT 1480
Db 1138 AGTTTCGTAATCTGCTCAATCAAGGAATCTATTGATCAATTTGATTAAGGTGAATACACAT 1197
Qy 1481 ATCGTCAGACAAATTTATGTCAATCCAAAGTGAGATACAGTTTATGCGCGGTTTAAACAG 1540
Db 1198 ATAGACAAACCGGTATATGTTAAATCCTAGTAGTACATCTGTAGTGGATCCTTATTAAAG 1257
Qy 1541 GTAATTTAAACCAAAATACGGATAGTAAATGCATTAATAGATACGAAATAACAGATTA 1600
Db 1258 GCGGTTCAATACCGGGTACCAATAGTAAATTAATAGATGACAGAAATACATCAATAA 1317
Qy 1601 AAGTATATAAGTAGATAATGTCAGCTGATTATCTGAAAGTTACTTTGTGAATCCAGAAA 1660
Db 1318 AAGTTTATTAAGTTAAGAAAAAGCGCATCTAACAGACAGCTATTATGTTGATCCATCAA 1377
Qy 1661 ACTTTGAGGATGTCATTAATAGTGTGAATATTACATTCGCCAAATCCAAATCAATATAAAG 1720
Db 1378 ATTATGAAGATGTTACAAAGTGACGTCAAAATTTACTTTCCCTGAAAAAGGAATATATCAAA 1437
Qy 1721 TAGAGTTTATACGCTGATCATCAAAATTAACACCGGTATATAGTAGTTGTTAATGCTC 1780
Db 1438 TCAACTTTAATACCTGAAGATGATCAAAATTAATCTCCATAGCTAGTTGTTATTAATGCTC 1497
Qy 1781 ATATTGATCCGAATAGCAAAAGGTGATTAGCTTTACGTTTCAACTTTATATGGGTATAACT 1840
Db 1498 ATGTTGATCCGAATAGTAATGGCAATTTGTTATTTAAGTTTCAACATTAATGTTTACGATT 1557
Qy 1841 CGAATATAATTTGGCGCTCTATGTCATGGGACACAGAGTAGCATTTTAATACCGATCAG 1900
Db 1558 CTAATTTTCACTAGAGTTTCTATGGCATGGGATAATGAAGTAGAATATCATGACGAAAGTG 1617
Qy 1901 GTTCTGTGACGGTATCGATAAAACAGTTGTTCTGTAACACCTGATGAGCTCGTGAAA 1960
Db 1618 GTAATGGTACGGTATAGATAAATCTGTAATACCGGATCAACCGGGGATTCGGAT---- 1673
Qy 1961 TTGAACCAATTTCCAGAGGATTCAGATTCGACCCAGGTTTCAGATTTCTGGCAGCGATTCTA 2020
Db 1674 -----AGTGTGATCATTTCCGATACGAGCGACAGCTCAGATAGCGACGAG 1719
Qy 2021 ATTACAGATAGCGGTTTCAGATTCGGGTAGTGTATCTACATCAGATAGTGGTTTCAGATTAC 2080
Db 1720 ACTCCGACAGCGATGCGGACTCCGATAGTGTAGATTTCCGACAGCGATGACACTCCG 1779

Qy 2081 CGAGTGTATTTCAGATTTCAGCAAGTGATTTCAGACTCAGCGAGTGATTTCAGATTTCAGCAAGCG 2140
Db 1780 ATAGTAGCGCAGATTTCGACAGCGATGCGAGATTCGAGATTCGAGACTCCGATAGCG 1839
Qy 2141 ATTCCGACTCAGCGCGGATTCGAGCTCAGACAAATGACTCGGATTCAGATAGGAGATTCTG 2200
Db 1840 ACGCAGACTCCGATAGTAGTGGGATTCGATAGTATGATGCGGATTCGATAGCGACGAG 1899
Qy 2201 ACTCAGACAGTGTACTCAGATTCCGACAGTGTACTCAGATTTCAGATAGCGATTCTGACTCAG 2260
Db 1900 ACTCAGATAGCGAGCGAGACTCCGACAGCGATGCGGACTCCGATAGTGTGCGGACTCCG 1959
Qy 2261 ACAGTGTACTCAGATTTCAGATAGCGATTTCAGATTTCAGATAGCGATTTCGACAGAGTG 2320
Db 1960 ATAGTGTAGTGTAGTTCGACAGCGATGCGACTCAGATAGTGTGCGGAGATTCGACAGCG 2019
Qy 2321 ATTCCGACTCAGACAGCGATTTCGACTCCGACAGTGTACTCCGACTCAGACAGCGATTTCAG 2380
Db 2020 ATGCGATTTCGATAGTGGGACTCCGACAGTGTGCGGACTCCGATAGTGTGCGGACTCCG 2079
Qy 2381 ATTCCGACAGTGTACTCCGACTCAGATAGCGATTTCGACTCAGATAGCGACTTCAGATTTCAG 2440
Db 2080 ACTCCGACAGCGATGCGGACTCCGACAGCGCGAGATTTCGATAGTGTGCGGACTTCAG 2139
Qy 2441 ACAGCGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGACAGAGTG 2500
Db 2140 ATAGCGATTTCGATAGTGGGACTCCGACAGATTTCGATAGTGTGCGGACTTCGACAGAGTG 2199
Qy 2501 ACTCAGATTTCGACAGTGTACTCCGATTTCAGATAGCGATTTCAGATTTCGACAGTGTACTCAG 2560
Db 2200 ACGCGATTTCGACAGTGTACTCCGATTTCGATAGTGTGCGGACTTCAGATAGTGTGCGGACT 2259
Qy 2561 ATTCCGACAGTGTACTCAGACTCAGACAGTGTACTCCGATTTCAGCGAGTGTACTCCGATTTCAG 2620
Db 2260 ATTCCGACAGCGATGCGAGATTTCGACAGCGATGCGAGATTTCGACAGCGCGGACTTCAG 2319
Qy 2621 ATAGTGTATTCCGACTCCGACAGTGTACTCCGATTTCAGATAGCGATTTCAGATTTCGAGATTTCG 2680
Db 2320 ATAGCGATTTCAGACTCCGACAGCGATGCGAGATTTCGATAGTGTGCGGACTTCGACAGAGTG 2379
Qy 2681 ACTCGGATTTCAGATAGCGATTTCGACTCAGATAGCGATTTCAGATTCAGACAGCGGATTTCAG 2740
Db 2380 ATGCGGACTCCGACAGTGTACTCCGATTTCGATAGCGCGAGATTTCGACAGCGAAATTCG 2439
Qy 2741 AATCAGACAGCGATTTCAGATTTCAGACAGCGACTTCAGACAGTG 2783
Db 2440 GTTCCGACAGCAATGTGGATCTCAGATAGTGTATTCAGAAATGA 2482

Search completed: November 5, 2004, 10:04:35

Job time : 14386 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:52:30 ; Search time 2.75536 Seconds
(without alignments)
3282.472 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425
Perfect score: 487
Sequence: 1 TLTPAYIDPENVKKTGNVT.....NVIAPVLTGNLKPNTDSNAL 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	100.0	933	2 S41539	fibrinogen-binding
2	422	86.7	989	2 D89852	fibrinogen-binding
3	113	23.2	1018	2 A32192	fibrinogen-binding
4	100.5	20.6	940	2 S19702	fibrinogen-binding
5	99.5	20.4	1092	2 T30214	fibrinogen-binding
6	88.5	18.2	1166	2 T28680	fibrinogen-binding
7	87.5	18.0	961	2 G90053	fibrinogen-binding
8	87	17.9	1141	2 E89824	fibrinogen-binding
9	84.5	17.4	1038	2 H90053	fibrinogen-binding
10	81.5	16.7	1315	2 T28679	fibrinogen-binding
11	78	16.0	202	2 G70380	probable membrane
12	76.5	15.7	1183	2 S63046	probable ribosomal
13	75.5	15.5	109	2 T38657	probable ribosomal
14	75.5	15.5	282	2 G64479	probable ribosomal
15	75.5	15.5	553	2 H64494	probable ribosomal
16	75.5	15.5	586	2 T51211	probable ribosomal
17	73.5	15.1	745	2 S13586	triacylglycerol li
18	72	14.8	3283	2 AC1018	large repetitive p
19	71.5	14.7	591	1 NUZ0F	glucose-6-phosphat
20	71.5	14.7	894	2 S54786	multifunctional be
21	71.5	14.7	1946	2 AE1449	hypothetical prote
22	71	14.6	877	2 F90070	Clumping factor B
23	70	14.4	391	2 E71125	probable aromatic-
24	70	14.4	415	2 B97085	DNA-methyltransfer
25	70	14.4	633	2 AD1116	internalin, probab
26	70	14.4	1959	2 AG1085	hypothetical prote
27	69.5	14.3	308	2 D90261	conserved hypothet
28	69.5	14.3	528	2 T41362	hypothetical prote
29	69.5	14.3	1305	2 D82923	DNA-directed RNA p

30	69	14.2	679	2 S43129	penicillin-binding
31	69	14.2	1872	2 S64851	probable membrane
32	68.5	14.1	154	2 B48230	T-cell receptor be
33	68.5	14.1	165	2 S31618	T-cell receptor be
34	68.5	14.1	995	2 S50358	hypothetical prote
35	67.5	13.9	329	2 A30760	1-phosphatidylino
36	67.5	13.9	345	2 F86368	hypothetical prote
37	67.5	13.9	549	2 C86692	peptide-binding pr
38	67.5	13.9	1132	2 S37932	hypothetical prote
39	67.5	13.9	1324	1 VGH59	E2 glycoprotein pr
40	67.5	13.9	1385	2 D89824	hypothetical prote
41	67	13.8	174	2 AF2441	hypothetical prote
42	67	13.8	230	2 T28262	ORF MSV100 probabl
43	67	13.8	450	2 B96561	hypothetical prote
44	67	13.8	512	2 AD1132	transcription regu
45	67	13.8	533	2 D91245	malate synthase A

ALIGNMENTS

RESULT 1

S41539
fibrinogen-binding protein - Staphylococcus aureus
N;Alternate names: clumping factor
C;Species: Staphylococcus aureus
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S41539; S36630
R;McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus
A;Reference number: S41539; MUID:94224142; PMID:8170386
A;Accession: S41539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-933 <WCD>
A;Cross-references: UNIPROT:Q53653; EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g39752

Query Match 100.0%; Score 487; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.2e-39;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TLTPAYIDPENVKKTGNVLTATIGSTTANKTLVDYKYGFYNSIKGTIDQIDKTN 60
Db	332	TLTPAYIDPENVKKTGNVLTATIGSTTANKTLVDYKYGFYNSIKGTIDQIDKTN 391
QY	61	NTYRQIYVNPNGDVIAPVLTGNLKPNTDSNAL 94
Db	392	NTYRQIYVNPNGDVIAPVLTGNLKPNTDSNAL 425

RESULT 2

D89852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D89852
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: D89852
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-989 <KUR>
A;Cross-references: UNIPROT:Q53653; GB:BA000018; PID:gl3700678; PIDN:BA041975.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: cifa

Query Match 86.7%; Score 422; DB 2; Length 989;

```

Best Local Similarity   84.9%; Pred. No. 2.9e-33;
Matches    73; Conservative      7; Mismatches     7; Indels       0; Gaps      0;

QY      2 LTPPAYIDPENVKKTGNVTLTATGGTSTANKTVLVDEKYKGKFNLSIKGTIDQIDDKTN 61
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      333 ITPAYIDPENVTKTGNVTITGIGINTASKTVLVDYEKYGFHNLSIKGTIDQIDKTN 392
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY      62 TYRQTYYNPSGDNIAPVLITGMNKPNWDSNAL 94
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      393 TYRQTYYNPSGDNVLPALTGNLIPTKSNAL 425
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

RESULT 3
A32192
fibronectin-binding protein - Staphylococcus aureus
C/Species: Staphylococcus aureus
C/Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997
C/Accession: A32192
R/Signaes, C.; Raucel, G.; Joensson, K.; Lindgren, P.E.; Anantharamaiah, G.M.; Hoeck, M.
Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989
A>Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylo-
A/Reference number: A32192; MUID:89098998; PMID:2521391
A/Accession: A32192
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1018 <SIG>
A/Cross-references: GB:J04151
C/Keywords: fibronectin binding

Query Match      23.2%; Score 113; DB 2; Length 1018;
Best Local Similarity 23.0%; Pred. No. 0.0059;
Matches    23; Conservative    24; Mismatches    43; Indels     2; Gaps     2;

QY      2 LTPPAYIDPENVKKTGNVTLTATGGTSTANKTVLVDEY-EKYKGKFNLSIKGTIDQIDKTN 60
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      301 LEINLFIDPRTVOTNGNQTTITLINEEQTSKLVDKYKGIGNYY-ANLNGSLIETFNKAN 359
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY      61 NTYRQTYYNPSGDNIAPVLITGMNLKPNTSDN 92
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      360 NRFSHVAFIKPNNGKTTSVTVTGTLMKGSQN 391
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

RESULT 4
S19702
fibronectin-binding protein B - Staphylococcus aureus
C/Species: Staphylococcus aureus
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S19702
R/Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A>Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus
A/Reference number: S19702; MUID:92111475; PMID:1837266
A/Accession: S19702
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-940 <OUE>
A/Cross-references: UNIPROT:Q53682; EMBL:X62992; NID:g49040; PIDN:CAA44726.1; PID:g58156
C/Keywords: fibronectin binding

Query Match      20.6%; Score 100.5; DB 2; Length 940;
Best Local Similarity 23.2%; Pred. No. 0.089;
Matches    23; Conservative    25; Mismatches    38; Indels    13; Gaps     3;

QY      2 LTPPAYIDPENVKKTGNVTLTATGGTSTANKTVLVDEY-----EKYGFNLSIKGTIDQI 56
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      273 LS.LNFLDPTTDTVTKGNQNVKEVLGETVSKIFNIQVLGGVRDNWG----VTANGRIDTL 328
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY      57 DKTNNTRYQTIYYNPSSGDNIAPVLITGNL----KNWTD 91
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db      329 NKVDGKFSPHYAKPPNNQSLSSTVTVGQVTKGNKPGVN 367
          :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

RESULT 5

```


A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: UNIPROT:Q99RD3; GB:BA000018; PID:gl3702452; PIDN:BAB43593.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb8

Query Match 18.0%; Score 87.5; DB 2; Length 961;
Best Local Similarity 22.5%; Pred. No. 1.7;
Matches 20; Conservative 21; Mismatches 39; Indels 9; Gaps 2;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKTVLVVDY-----EKYKGFYNLSIKGTIDQI 56
Db 276 LNLNLFIDFTTQKGSQKVEVLGQNKVSKFEIDIKYLDGVKDRMG---VTWNGRIDTL 331

QY 57 DKTNNTYRQTIYVNPNGDNVIAPVLTGNL 85
Db 332 NKEEGKFSHFAYYKPNQSLTSVTVTGQV 360

RESULT 8
E89824
hypothetical protein sdrE [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89824
R:Kuroda M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E89824
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1141 <KUR>
A:Cross-references: UNIPROT:Q99W46; GB:BA000018; PID:gl3700455; PIDN:BAB41752.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrE

Query Match 17.9%; Score 87; DB 2; Length 1141;
Best Local Similarity 34.7%; Pred. No. 2.4;
Matches 25; Conservative 14; Mismatches 29; Indels 4; Gaps 2;

QY 2 LTMPAYIDPENV--KKTGNVLTATIGSTTANKTVLVVDYKYGKGFYNLSIKGTIDQIKT 59
Db 382 LTLYSYIDKQAVPNETSLSLTTATAGKTSQN--VSDYQDPVMVHGDSNIQSFTKLDEN 439

QY 60 NNTYRQTIYVNP 71
Db 440 KQTIEQQIYVNP 451

RESULT 9
H90053
hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H90053
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: UNIPROT:Q99RD2; GB:BA000018; PID:gl3702453; PIDN:BAB43594.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:

A:Gene: fnb

Query Match 17.4%; Score 84.5; DB 2; Length 1038;
Best Local Similarity 24.7%; Pred. No. 3.7;
Matches 21; Conservative 17; Mismatches 46; Indels 1; Gaps 1;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKTVLVVDYKYGKGFYNLSIKGTIDQIDKTN 61
Db 295 LNLNLFIDPKTVQSGEQKITSKLNGEETKTEKIPVYVNEGVSNYTNVNGSIETFKESN 354

QY 62 TYRQTIYVNPNGDNVIAPV-LTGNL 85
Db 355 KETHIAYIKPMNGNQSNSTVSTGTL 379

RESULT 10
T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28679
R:Josselson, E.; McCrea, K.; Ni Bhdhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.;
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of Sta
A:Reference number: Z20510; MUID:99098700; PMID:9884231
A:Accession: T28679
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1315 <JOS>
A:Cross-references: UNIPROT:O86488; EMBL:AJ005646; NID:el318791; PID:el318792; PIDN:CAA06
C:Genetics:
A:Gene: sdrD

Query Match 16.7%; Score 81.5; DB 2; Length 1315;
Best Local Similarity 31.2%; Pred. No. 9.8;
Matches 24; Conservative 12; Mismatches 34; Indels 7; Gaps 3;

QY 7 YIDPENVKKTGN-VLTATIGSTTANKTVLVVDYKYGKGFYNLSIKG---TIDQIDKTN 61
Db 361 YMDADTFPVKNDVFNVTGNTTKTANIQPDYVNEKNSIGSAFTETVSHVGNKEN 420

QY 62 T--YRQTIYVNPNGDNV 76
Db 421 PGYKQTIYVNPSENLS 437

RESULT 11
G70380
DNA polymerase III epsilon subunit - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: G70380
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: G70380
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-202 <AQF>
A:Cross-references: UNIPROT:O67074; GB:AE000714; NID:g2983446; PIDN:AA07031.1; PID:g2983
A:Experimental source: strain VF5
C:Genetics:
A:Gene: dnaQ

Query Match 16.0%; Score 78; DB 2; Length 202;
Best Local Similarity 31.6%; Pred. No. 2.2;
Matches 25; Conservative 10; Mismatches 32; Indels 12; Gaps 3;

QY 1 TLTMPAYIDPENVKKTGNVLTATIGSTTANKTVLVVDYKYGKGFYNLSIKGTI 53
Db 55 TLVYPGYFIPERIKKLTGITNAMLVGQPTIE-VLPEFLEFVGDNVVGHFVEQDIK--- 110

QY 54 DQIDKTNNTYRQTIYVNP 72
Db 111 -FINKYTKQYKGRNPS 128

RESULT 12
S63046
Probable membrane protein YNL106C - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N2160
C:Species: Saccharomyces cerevisiae
C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C:Accession: S63046; S63047; S58246; S67353; S63948
R:Saiz, J.E.; Buitrago, M.J.; Soler-Mira, A.; Del Rey, F.; Revuelta, J.L.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63037
A:Accession: S63046
A:Molecule type: DNA
A:Residues: 1-1102 <SAI>
A:Cross-references: UNIPROT:P50942; EMBL:Z71382; MIPS:YNL106C
A:Experimental source: strain S288C
R:De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63047
A:Accession: S63047
A:Molecule type: DNA
A:Residues: 880-1183 <DEA>
A:Cross-references: EMBL:Z71382; MIPS:YNL106C
A:Experimental source: strain S288C
R:Saiz, J.E.; Buitrago, M.J.; Soler, A.; del Rey, F.; Revuelta, J.L.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of a 21.3 kb fragment from the left arm of yeast chromosome XI
A:Reference number: S58246
A:Accession: S58246
A:Molecule type: DNA
A:Residues: 1-1102 <SAW>
A:Cross-references: EMBL:Z50161; NID:G929846; PID:G929847
R:d'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.
submitted to the EMBL Data Library, February 1996
A:Description: The DNA sequence of cosmid 14-13b from chromosome XIV of Saccharomyces cerevisiae
A:Reference number: S67327
A:Accession: S67353
A:Molecule type: DNA
A:Residues: 880-1183 <DAN>
A:Cross-references: EMBL:Z69382; NID:gl183941; PID:gl183968
R:Saiz, J.E.; Buitrago, M.J.; Soler-Mira, A.; del Rey, F.; Revuelta, J.L.
Feast 12, 403-409, 1996
A:Title: The sequence of a 21.3 kb DNA fragment from the left arm of yeast chromosome XI
A:Reference number: S63948; MUID:96267765; PMID:8701612
A:Accession: S63948
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 662-887 <SAF>
A:Cross-references: EMBL:Z50161
C:Genetics:
A:Gene: SGD:INP52
A:Cross-references: SGD:S0005050; MIPS:YNL106C
A:Map position: 14L
C:Keywords: transmembrane protein
F:62-78/Domain: transmembrane #status predicted <TM1>
F:667-683/Domain: transmembrane #status predicted <TM2>

Query Match 15.7%; Score 76.5; DB 2; Length 1183;
Best Local Similarity 33.3%; Pred. No. 27;
Matches 21; Conservative 5; Mismatches 32; Indels 5; Gaps 2;

QY 12 NVKKTGNVLTATGIGSTTANKTVLDVDEYKYGK-----FYNLSIKGTIDDKTNNTYRQTI 67
Db 690 NIKEVGGSTKTKTGFGGITGNKGAIVRFDYGATSCFVNTLSAGASNIDRRNDY--NNI 748

QY 68 YVN 70
Db 749 YRN 751

RESULT 13

T38557

Probable ribosomal protein, mitochondrial - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38657
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21804
A:Accession: T38657
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-109 <MUR>

A:Cross-references: UNIPROT:Q9UT82; EMBL:AL109739; NID:e1534774; PID:CAB52271.1; GSPDB:C

A:Experimental source: strain 972h-; cosmid c343

C:Genetics:

A:Gene: SPDB:SPAC343.08c

A:Map position: 1

A:Genome: nuclear

A:Introns: 20/3; 104/2

C:Keywords: mitochondrion

RESULT 15

H64494
hypothetical protein M01561 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: H64494
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
R.; Raich, C.I.; Overbeek, R.; Kirtress, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
Ison, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8668087
A:Accession: H64494
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-553 <BUL>
A:Cross-references: UNIPROT:Q58956; GB:U67596; GB:I77117; NID:G2826430; PIDN:AAB99582.1;
C:Genetics:
A:Map position: REV1538169-1536508

```

Query Match      15.5%; Score 75.5; DB 2; Length 553;
Best Local Similarity 36.4%; Pred. No. 13;
Matches 24; Conservative 7; Mismatches 22; Indels 13; Gaps 5;

Qy      12 NVKKTGNVTATGIGSTTANKTVLDYEKY--GKFYNLSIKGIDQIDQKNTNVRQ---- 65
      ||| ||| : ||| : ||| : ||| : ||| : ||| :
Db 346 NIKDVGN--LLPGVSKTV--SFVVDVDKVASAKHYKLPIE--ISYLDTANNKYKTEKFI 398

Qy      66 TIYNVP 71
      ||| |
Db 399 DIYVKP 404

```

Search completed: November 3, 2004, 19:04:08
Job time : 3.75536 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:52:30 ; Search time 15,4769 Seconds
(without alignments)
3282.472 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VGTLLIGFGLSSKEADASEN.....NGSGSGDGDKPWPEQPD 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2702	100.0	933	2 S41539	fibrinogen-binding
2	2466	91.3	989	2 D89852	fibrinogen-binding
3	520	19.2	961	2 G90053	hypothetical prote
4	515	19.1	940	2 S19702	fibrinogen-binding
5	494.5	18.3	1166	2 T28680	fibrinogen-binding
6	494	18.3	1141	2 B89824	hypothetical prote
7	472	17.5	1092	2 T30214	fibrinogen-binding
8	467	17.3	877	2 F90070	clumping factor B
9	452.5	16.7	1018	2 A32192	fibrinogen-binding
10	441.5	16.3	1038	2 H90053	hypothetical prote
11	398	14.7	1315	2 T28679	fibrinogen-binding
12	379	14.0	1385	2 D89824	hypothetical prote
13	356	13.2	953	2 C89824	hypothetical prote
14	249	9.2	797	1 VGBX1	glycoprotein x pre
15	248.5	9.2	1072	2 A86827	hypothetical prote
16	236	8.7	995	2 S03558	hypothetical prote
17	234.5	8.7	2271	2 F90073	hypothetical prote
18	229	8.5	750	2 T42613	probable envelope
19	228.5	8.5	3507	2 T34513	hypothetical prote
20	226.5	8.4	1777	2 T34369	hypothetical prote
21	225.5	8.3	1459	2 T45463	membrane glycoprot
22	224	8.3	1459	2 T32271	hypothetical prote
23	223.5	8.3	833	2 E90577	lipoprotein vsaC
24	223.5	8.3	967	2 S66852	hypothetical prote
25	223.5	8.3	1161	2 S57180	probable membrane
26	223	8.3	866	2 T45462	membrane glycoprot
27	223	8.3	1125	2 E90598	membrane nuclease,
28	221.5	8.2	1367	1 S48478	glucan 1,4-alpha-g
29	220	8.1	1758	2 S57015	probable purine nu

30	218.5	8.1	2232	2 T34434	hypothetical prote
31	217	8.0	725	2 A41258	a-egglutinin core
32	217	8.0	792	2 S70305	hypothetical prote
33	216	8.0	796	2 T21460	hypothetical prote
34	214	7.9	1063	2 D86731	hypothetical prote
35	212	7.8	1107	2 AC0976	probable autotrans
36	209.5	7.8	971	2 B90835	probable tail fibre
37	209.5	7.8	973	2 C85693	probable membrane
38	207	7.7	528	2 I47141	gastric mucin (clo
39	207	7.7	551	2 S64314	probable membrane
40	207	7.7	1275	2 T33369	hypothetical prote
41	206.5	7.6	1229	2 T25697	hypothetical prote
42	205	7.6	888	2 T46726	secreted acid phos
43	205	7.6	1104	2 S59310	probable membrane
44	204	7.5	1169	2 S38181	floculation prote
45	203	7.5	1829	2 T24583	hypothetical prote

ALIGNMENTS

RESULT 1

S41539
fibrinogen-binding protein - Staphylococcus aureus
N;Alternate names: clumping factor
C;Species: Staphylococcus aureus
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S41539; S36630
R;McDevitt, D.; Francois, P.; Vaudaux, P.; Poster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staph
A;Reference number: S41539; MUID:94224142; PMID:8170386
A;Accession: S41539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-933 <WCD>
A;Cross-references: UNIPROT:Q53653; EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g39752

Query Match	100.0.0%;	Score	2702;	DB 2;	Length	933;			
Best Local Similarity	100.0.0%;	Pred. No.	4.4e-125;	Mismatches	0;	Indels	0;	Gaps	0;
Matches	528;	Conservative	0;						
Qy	1	VGTLLIGFGLSSKEADASENSVTQSDSASNEKSDSSVSAAAPKTDNTVSDTKTSNT	60						
Db	23	VGTLLIGFGLSSKEADASENSVTQSDSASNEKSDSSVSAAAPKTDNTVSDTKTSNT	82						
Qy	61	NGGETSVAQNPAQOETTQSSSTNATTEPVTGATTTTNTQANTPATTOSSNTNABELV	120						
Db	83	NGGETSVAQNPAQOETTQSSSTNATTEPVTGATTTTNTQANTPATTOSSNTNABELV	142						
Qy	121	NOTSNTEFTNDTNTVSSVNSPQNSTNAENVSTTQDTSTATPSNNESAPOSTDASNKDVV	180						
Db	143	NOTSNTEFTNDTNTVSSVNSPQNSTNAENVSTTQDTSTATPSNNESAPOSTDASNKDVV	202						
Qy	181	NOAVNTSAPRMRAFSLAAVAADAPAAAGTDTNLTNTVTGIDSGTTTYPHQAQYKVLNYG	240						
Db	203	NOAVNTSAPRMRAFSLAAVAADAPAAAGTDTNLTNTVTGIDSGTTTYPHQAQYKVLNYG	262						
Qy	241	FSVPNSAVGDTFKIIVPKELNNGVTSTAKVPPIMAGDQVLANGVIDSGNVIYFTDY	300						
Db	263	FSVPNSAVGDTFKIIVPKELNNGVTSTAKVPPIMAGDQVLANGVIDSGNVIYFTDY	322						
Qy	301	VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGTGSTANKTGLVDYKGFYKFNLSIKG	360						
Db	323	VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGTGSTANKTGLVDYKGFYKFNLSIKG	382						
Qy	361	TIDQIDKTNNTYQTIIYVNPSPGDNVIAVLTLGNLKNPNTDSNALIDQONTISIKYKVDNAA	420						
Db	383	TIDQIDKTNNTYQTIIYVNPSPGDNVIAVLTLGNLKNPNTDSNALIDQONTISIKYKVDNAA	442						
Qy	421	DLSESYFVNPNEDVTNSVNIITFPNPQYKVFNTPDDQITTPYIVVNGHIDPNKSGD	480						
Db	443	DLSESYFVNPNEDVTNSVNIITFPNPQYKVFNTPDDQITTPYIVVNGHIDPNKSGD	502						

QY 481 LALRSTLYGNSNIIWRSMNDNEVAFNNGSGDGDIDKPVVPEOPDE 528
|||||
Db 503 LALRSTLYGNSNIIWRSMNDNEVAFNNGSGDGDIDKPVVPEOPDE 550
|||||
RESULT 2
D89852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
A:Accession: D89852
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.;
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: D89852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-989 <KUR>
A:Cross-references: UNIPROT:Q99VJ4; GB:BA000018; PID:g13700678; PIDN:BAB41975.1; GSPDB:C
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfa

Query Match 91.3%; Score 2466; DB 2; Length 989;
Best Local Similarity 91.1%; Pred. No. 1.7e-113;
Matches 481; Conservative 19; Mismatches 28; Indels 0; Gaps 0;
QY 1 VGTLLGFGLLSKEADASENSVTQSDSASNESKNSDSSVSAAPKTDNTVSDTKTSNT 60
Db 23 VGTLLGFGLLSKEADASENSVTQSDSASNESKNSDSSVSAAPKTDNTVSDTKTSNT 82
QY 61 NNGETSAQNPAQOBTQSSNTATTEPTVTEATTTTQANTPATQSSNTNAEELV 120
Db 83 NNGETSAQNPAQOBTQSSNTATTEPTVTEATTTTQANTPATQSSNTNAEELV 142
QY 121 NOTSNETTNTVSSVNSPQNSNAENVTQDTSFEAPSPNNEAPQSTDAENKDVV 180
Db 143 NOTSNETTNTVSSVNSPQNSNAENVTQDTSFEAPSPNNEAPQSTDAENKDVV 202
QY 181 NOAVNTSAPRMRAPSLAAVAADAPAGTDTITNQLTNVTGIDSGTIVPHQAGYVKNY 240
Db 203 SOAVNPSPTMRAPSLAAVAADAPAGTDTITNQLTDVKTIDSGTIVPHQAGYVKNY 262
QY 241 FSVNSAVKGTDFKTIPTPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
Db 263 FSVNSAVKGTDFKTIPTPKELNLTGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVKATITMPAYIDPENVKITGNVTLATGISTTANKTLDVDEKYGFYNLSIKG 360
Db 323 VDKNENVTANITMPAYIDPENVKITGNVTLATGISTTANKTLDVDEKYGFYNLSIKG 382
QY 361 TIDQIDKTNTRYQTIYVNSPQNSVIAPLVTCNLKPNPDSNALIDQOQNTSIKVKVDNAA 420
Db 383 TIDQIDKTNTRYQTIYVNSPQNSVIAPLVTCNLKPNPDSNALIDKNTDVKYVDNAN 442
QY 421 DLSSEYFVNPNFEDVNSVNIITPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 480
Db 443 DLSSEYFVNPNFEDVNSVNIITPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 502
QY 481 LALRSTLYGNSNIIWRSMNDNEVAFNNGSGDGDIDKPVVPEOPDE 528
Db 503 LALRSTLYGNSNIIWRSMNDNEVAFNNGSGDGDIDKPVVPEOPDE 550

RESULT 3
G90053
hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: G90053
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.;
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: G90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: UNIPROT:Q99RD3; GB:BA000018; PID:g13702452; PIDN:BAB43593.1; GSPDB:C
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnbB

Query Match 19.2%; Score 520; DB 2; Length 961;
Best Local Similarity 26.8%; Pred. No. 3.9e-18;
Matches 149; Conservative 89; Mismatches 201; Indels 118; Gaps 15;
QY 1 VGTLLGFGLLSKEADASENSVTQSDSASNESKNSDSSVSAAPKTDNTVSDTKTS 57
Db 21 LGTIVVVGQKEKAAASEQNNTTVESGSGATSKA-----SETQTT 63
QY 58 SNTNNGETSAQNPAQOBTQSSNTATTEPTVTEATTTTQANTPATQSSNTNAE 117
Db 64 TNNVNTIDETQSYSATSTEQPSKSTQVTEAPTVOAPKVETEMKS----- 110
QY 118 ELVNQTSNETTNTVSSVNSPQNSNAENVTQDTSFEAPSPNNEAPQSTDA 174
Db 111 -----QEDLPSEKVDKETTGTQVDIAQPSN----- 136
QY 175 SNKDVVNCVNTSAPRMR-----AFSLAAVAADAPAGTDTITNQLTNVTGID---SG 224
Db 137 -----VSEIKPRMKRSADVTAVSEKEVAEAKATGTDVTKNKEVTESSLEGHNKDS 187
QY 225 TTVPHQAGYVKNYGFSPNSAVKGTDFKTIPTPKELNLTGVTSTAKVPPIMAG--DOVL 282
Db 188 NVNPHNAQRVTLAYKWKFGSGIKAGDYDFTLSDNVETHGISTLRKYPEIKSSDEKVM 247
QY 283 ANGVISDGNVIYTFDYVNTKDDVKATITMPAYIDPENVKITGNVTLATGISTTANKT 342
Db 248 ANGVINERTIRYTFDYINNKDILTALNELNLNLFIDPTTVTKQSGKVEVTLGQNKVSK 307
QY 343 VLVDY-----EKYGFYNLSIKGTIDQIDKTNTRYQTIYVNSPQNSVIAPLVTCNLKPN 397
Db 308 FDIKYLGVKDRMG-----VTVNGRIDTLNKEEGKFSHPAYVYKPNQSTSVTVTGQVTS 363
QY 398 TDSNALIDQOQNTSIKVKVDNAA DLSESYFV---NPNFEDVNSVNIITPNPNQYKVEF 454
Db 364 YKQSA---NNPTVKVYKHIGSDLAESVYAKLDDTSKFEDVTEKVNLSYTSNGGYTLNL 419
QY 455 NTPDQITTPYIVVNGHIDPNSKGDALRSLTLGVNSNIIR-----SMWDNEVA 506
Db 420 GDLDN--SKDVIKYEGEYDQNAK-DLNFRTHLGSGHYKYPYYPYYPYVQLTWNGVA 476
QY 507 FNNGSGSGDGDIDKPVV 523
Db 477 FYSNAKGDGDKNDP 493

RESULT 4
S19702
fibronectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
R;Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus
A:Reference number: S19702; PMID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary

Matches	172	Conservative	91	Mismatches	215	Indels	131	Gaps	27
QY	1	VGTLIGFGLLSKKA	DASENSVTQSD	SASNEKSN	DSSVSAAPKTD	TNNVSDTKTS	SN	60	

[illegible]

Qy	61	NNGETSVAQNPAQOETTQSSSTNATTEPTVGEARITTTTNNQANTPATQSSNTNNAE--EL	1119
Db	85	ENDST---NPIKKEITNDISOPEAKEEST---TSSTQQOQNNVTATETTKPQIEKEN	135
Qy	120	VNQTSNETTFNDTIV-----SSVNSPQSTNAENVUS---T	152
Db	136	VKPSTDKTATEDTSVILLEEKAPNVNNDVTTKPSSTBIQTKPTTFQBSTNJENSOQPT	195
Qy	153	TQDTSTEATPSNSESAP-----QSTDASNKDVQNVQAVNTSAPR	190
Db	196	PSKYDNQVTDATNPKPVPVNSKEELKNPEKIKELVRNDNNTDRSTKPVATPTSVAPKR	255
Qy	191	MRAFSLAAVAADAPAAAGTDIINQJLT---NVTVGIDSGTIVVPHQAGVYKVLNUGFSVNS	246
Db	256	LNAKMRFAVAPAAAVASNNVNDLITVTKQTIKVGDKDNVAAAHGDKDIEYDTEFTIDNK	315
Qy	247	AVKGGTTFKI-----TVPKELNNGVTSHTAKVPPIMAGD---QVLANGVID--SDGNVITTF	297
Db	316	VKGGBTMTINVDKNVIPSGL-----TDKNDDIDITDPSGEBVIARAGTFDKATKQITVTF	368
Qy	298	TDVYNTKDDVKATLTMPIAYIDPENV--KKTGNVTIATGIGSTTANKTVLVDYEKYCKFVN	355
Db	369	TDYVDKYEDIKARLTLYSIDQKQAVPNETSJLNFATAGKETSON--VSVDYQDPMWCHD	426
Qy	356	LSIKGTIDQIDKNTNVTQTTIYNVP---SGDNVIAPVL-----TGNIKPNTDSNALIDQ	406
Db	427	SNIQSIPTKLDENKQTBQIYYVNPPLKKTATNTKVDIAGSQVDDYGNIKLGNGS--TIID-	484
Qy	407	QNTSIKVKYKVDNAADLSESYEV--NPENFEDVTNSVNIITFPNPQYKVBFTPD--DQITTF	464
Db	485	QNTSEIKVKYKVPNOQLPQSNRIYDFSQYEDVTSQ-----FDKNKSPSNVATLDFGDINSA	540
Qy	465	YIVVNGHIDPNKXGDL-----ALRST--LYGYSNIIWRSMWNVEAFNNGSGGQI	517
Db	541	YIIKVKSYTPTSOGELDIAQTSMTRTDKYGY-----YNTAGYSNFIVTSNDTGGDGT	595
Qy	518	DKP	520
Db	596	VKP	598

RESULT 7
T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.
Infect. Immun. 66, 2666-2673, 1998
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: Z20781; MUID:38261511; PMID:9596732
A:Accession: T30214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <N1>
A:Cross-references: UNIPROT:O70022; EMBL:Y17116; NID:e1296734; PDB:e1296735; PDB:CAA766

	Query Match	17.5%;	Score 472;	DB 2;	Length 1092;	
	Best Local Similarity	26.5%;	Pred. No. le-15;			
	Matches	157;	Conservative 103;	Mismatches 238;	Indels 94;	Gaps 25
Qy	1	VGTLLIGGLLSSKKADASENSV-----TQSDSASNESKSNDSVVSAAPKTDD	48			
Dd	36	IGATLLFG-LGHNEAKAEENSVDVKOSNTDDESDNDQSDEEKDVNNNQINTDD	94			
Qy	49	TNV-----SDTKTSNNTN-----NGETSVAGNP-----AQOETTGOSSSTN	83			
Dd	95	NNQLIKKEENNYDGIEKRSEDRTESTTNVDENEATFLQTPODNTHLTREEVKESVSVE	154			
Qy	84	ATTETPTVGBATTTTINQAANTPATQS-SNNIAEELVN---QTSNETTFNDTNV----	135			
Dd	155	SSNSIDTAOPSPHTTINRESVOTSDNVESHVSDPFANSKITKESNTSGKEENTIEOPN	214			

```

136 Qy -----SSVNSPQNSTNAENVSTTQDTS TEATPSN---NESAPQSTDAENKOVVNQAVNTS 187
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
215 Db KVKEDSTTSOPSGYTNIDEKISNQDELIN-LF1NEYENKARPLSTSAQPSIKRTVTN-- 271
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
188 Qy APRMRAFLAAVAADAPAAGGTDITNOLTNVTVGI-DSGTTVVYPHOAGYVKLNYSFGSPVNS 246
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
272 Db -----QLAAOQGSNNVHLIKVTDQ--SITEGYDDSEGVKAHDAENLIVDTVEVDKK 322
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
247 Qy AVKGDTFEKITVPKELNLSNGVTS TAKVPPIMAGD-QVLANGVIDS-DGNVIYFTFDYVNTK 304
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
323 Db VKSGDTWTVDIDKNVTSPDLTSDTFPIKDNSGEIIATGYDNKKKQITVFTFDYVVKY 382
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
305 Qy DVVKATLTTPAYIDPENPVKKTG--NVVLATGIGSTTANKTLVLDYEYKGFYNLSIKGT 361
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
383 Db ENIKAHLKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKIITVEYQRPNENRNTANLQSM 440
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
362 Qy IDQIDKTNNTYRQTIYVNPBGDNVIA PVLTKNLKPNTDSNALIDQONTSIKYKVVDNAAD 421
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
441 Db FTNIIDTKNHTVEQTYIYNLP--RYSAKETNVI NISGNDGEGSTIIDDSITIIKYKVGDNQN 498
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
422 Qy LSESIFYV-NPENPFEDTVNSVNITFPENPOYKVEFNTPDQTTPYIVVNGHIDPNSGKD 480
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
499 Db LPDSNRIYDVXEYVEDVNDYYAQLGNNDVNINFG----NIDSPIIIVKISKYDEN-KDD 553
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 Qy -----LALRSTLYGNSNI IWRSMGMDNEVAFNNGSGDGIDKPVVPEQ 525
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
554 Db YTTIQQIVTWGTTINEYTG--FRTASVDNTIAFSTSSGGQG-DLP--PEK 600
      ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: F90070
C: Mizutani, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, R.; Kuroda, M.; Ohta, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki, M. A.; Mizutani-Tai, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; PMID:21311952; PMID:11418146
A/Accession: F90070
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <CUR>
A/Cross-references: UNIPROT_Q9R907; GB:BA000018; PID:gl3702588; PIDN:BAB43728.1
A/Experimental source: strain N315
C:Genetics:
A:Gene: clfB

```

Query Match	17.3%	Score 467;	DB 2;	Length 877;
Best local Similarity	25.8%;	Pred. No. 1.4e-15;		
Matches 150;	Conservative 109;	Mismatches 204;	Indels 118;	Gaps 27;
Qy	1	VGTLLGFLGLSSKEADASE--NSVTOGDSASNESKGNSSSVGAAPKTDITDINVSDDTKSS	58	
Db	29	VGATILFG-ICNHQAQASEQSDNTTOS--SKNNSASDESKNMLETQLNTANDTSDISA	86	
Qy	59	NTNNGETSVAGNPAQGETTOSSTN-ATTEETPTVGTGEATTTTNTQANTPATTOSSVNAE	117	
Db	87	NTSNANDSTTKEMSTQTSNTITTEPASTNETP---OPTAIKNQAT--AAKQDQDTVPQ	140	
Qy	118	ELVNQTSNETFTDINTVSSVNSPQNSSTAEVNSTQDTSSTATSPNNEAPOSTASNK	177	
Db	141	EANSQVDNKTT-NDANSI-----ATNSE----LKNSQTLQLP--QSSPQT-----	178	
Qy	178	DVYNQAVNTSAPMRAPFSLAAVAADAP-----AAGTDITNQLTNTVGTDSSTVYPHQ	231	
Db	179	--TSNAGGTSKPSVRTEAVSSSLAVBPVNVAAQAKGVNNDKVTASNFKLEK--TTDPNQ	235	
Qy	232	AGVVKLYNGFVSNSAVKGGTFTKITVPKELNLTNGVTSTAKVPPIMAGDQVLAAGVID---	288	
Db	236	SGNTEMAANTFTVDKKVSGDYF-----TAKLP-----DSLTGNGDVPDYSN	275	

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-953 <KUR>
A;Cross-references: UNIPROT:Q99W48; GB:BA000018; PID:gl3700453; PIDN:BA41750.1; GSPDB:G
A;Experimental source: strain N315
A;Gene: sdrC

Query Match 13.2%; Score 356; DB 2; Length 953;
Best Local Similarity 25.8%; Pred. No. 4.2e-10;
Matches 140; Conservative 81; Mismatches 217; Indels 104; Gaps 23;
QY 1 VGTLLIGFLLSSKEADASENSVTQSDASNESKNSDSSVSAAAPKTDITNVSDTKTSNT 60
DB 35 VGTLLIFG-LSGHEAKAAEHNGELNOSKNETTA--PSENKTYEKVDSRQLKD----- 84
QY 61 NGETSVAQNPAAQOETTQSSSTNATBETPTVGEATTTTINQANTPATTOSSNNAABELV 120
DB 85 -NTQTATADQP--KVTMSDS--ATVKE-----TSSNMOSP-----QNATASQST 123
QY 121 NCTSNETTEND--TNTVSSVNSPQNSTNAENVSTTQDTSTTEATPSNNEAPOSTDAENKD 178
DB 124 TQTSNVTT--NDKSTTYSNETDKNLTQAKNVST-----PKTTTIKORA 167
QY 179 VVQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQL--TNTVTVGIDSG-----TTVY 228
DB 168 LNEMAVNTVA-----APOQGTNVNDKVHFTINIDALDKGVNKTGTGTEFW 213
QY 229 PHQAGVVKLVNYSFVNSAVKGTFTKTVKPELNLNGVTSTAKVPPIM-AGDOVLANGVI 287
DB 214 ATSSDVLKLVKANTIDDSVKEGDTFTFKYQGYFRPGSVRLPSQTONLYNAQNIKAGIY 273
QY 288 DSDGN-VIYFTDYVNTKDVKATLTPAYIDEN--VKKTG---NVTLATGSGSTANK 341
DB 274 DSKNTTITFTNVDOQYTNVSGSFEOVAFKRENATDKYAKMEVTL---GNDTYSK 329
QY 342 TLVDYKYGKFNLSIKGTIDQIDKTNNTYRTIYVNPNGDNVIAPLVLTGNT---KPNT 398
DB 330 DVIVDYGNGKQQLISSTNYINNEDLSRN--MTVYVNPQPKTYKETFVNTLGTGYKFN 386
QY 399 DSNALIDQOQNTSIKYKVDNAADLSESYFVNPFEDVTSVNTIYFPNPNQ-YKVENFTP 457
DB 387 DAK-----NFKIYEVTDQNFQVDSFTPDTSKLKDVQGFQDIYSNDKNKTATVDLLG 438
QY 458 DDQITTPYIVVNGHIDPNKGLALRSTLYGYSNLIWMSMDNEVAFNGSGSDGI 517
DB 439 QSSSDKQIILIQQVAYPNSSDTNGKIDYITLETQNG-----KSSWNSYSNVNNGSSTANGD 493
QY 518 DK 519
DB 494 QK 495

RESULT 14
VGBEX1
C;Species: equine herpesvirus 1 (strain Ab4p)
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: H36802
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: H36802
A;Molecule type: DNA
A;Residues: 1-797 <TEL>
A;Cross-references: UNIPROT:P28968; GB:M86664; NID:g330791; PIDN:AAB02506.1; PID:g330864
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566; PMID:1318606
A;Contents: annotation; possible protein-coding frames

A;Note: neither amino acid nor nucleotide sequence is given
C;Genetics:
A;Gene: 71

C;Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homolog
C;Keywords: glycoprotein; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-797/Product: glycoprotein X #status predicted <MAT>
F;23-465/Region: serine/threonine-rich
F;489-797/Domain: equine herpesvirus 1 glycoprotein homolog <EHG>
F;766-790/Domain: transmembrane #status predicted <TMN>
F;590/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.2%; Score 249; DB 1; Length 797;
Best Local Similarity 21.8%; Pred. No. 5.8e-05;
Matches 114; Conservative 80; Mismatches 238; Indels 90; Gaps 17;
QY 3 TLIGFLLSSKEADASENSVTQSDASNESKNSDSS-----VSAAPKTDITNVSDTK 55
DB 16 SIYAIGSTVTTTETTTSSSTSGSGQSTSGTSSSSPTTSPPTSSSPPTSTHTSSPSS 75
QY 56 TSNNTNGETSVQNPAAQOETTQSSSTNATBETPTVGEATTT-----TNQANTPA 107
DB 76 TSQSSSTATSSASPTASSTSIPTSTSTETTTPTTASITTTPTTTTAAATTAATTA 135
QY 108 TTQSSNTNAB--ELVNOTSNETTENDTNTVSSVNSPQNSTNAENVSTTQDTSTTEATPSN 164
DB 136 VTTAASATSETTATATATSTPTTTPTSTTTTATTTVPTTA---STTTDTTAAATTA 192
QY 165 NESAPQSTDSNKKDVNVQAVNTSAPRMRAFSLAAVAADAPAAAGTD-----ITNLTN 216
DB 193 ATTAAATTTAA--TTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 249
QY 217 VTGIDSGITVPHQAGVVKLVNYSFVNSAVKGTFTKTVKPELNLNGVTSTAKVPPIM 276
DB 250 TAAATTTAATTTGSGTSTT---GASTSTPSASTSATPTSTSTSAATSTPTPTS 306
QY 277 AGDQVLANGVIDSGNVIVFTDYVNTKDVKATLTPAYIDENPVNKKTKGNVTLATGIGS 336
DB 307 AATS--AESTTEAP-----TSTFTTDTTPSEAT---TATTSPESTVSASTSATTTAF 356
QY 337 TTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYR--QTIYVNPNGDNVIAPLVLTG-- 392
DB 357 TTESHT-----SPDSSSTGSTSTAEPSSTFTLTPTSTATPTSDQFTGSS 398
QY 393 -----NLKPNTDNSNALIDQOQNTSIKYKVDNAADLSESYFVNPFEDVTSVNTIYFPNPNQ 440
DB 399 ASTESDTSSTVPTTGTESITESSTT-----EASTNLGSGSTYESTEALETDPDGNNTS 452
QY 441 NITFPNPQYKVEF---NTPDDQITTPYIVVNGHIDPNK 478
DB 453 GNTTPSPRPFPSPFADTQQTDPNGVSTQH-TTINDHTTANAQ 493

RESULT 15
A86827
C;Species: Lactococcus lactis subsp. lactis (strain IL1403)
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86827
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A;Reference number: A86825; MUID:21235186; PMID:11337471
A;Accession: A86827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1072 <STO>
A;Cross-references: UNIPROT:Q9CF64; GB:AE005176; PID:gl2724625; PIDN:BAK05715.1; GSPDB:G
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yqfG

Query Match 9.2%; Score 248.5; DB 2; Length 1072;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:52:30 ; Search time 27.3484 Seconds
(without alignments)
3282.472 Million cell updates/sec

Title: US-09-679-643-2

Perfect score: 4725

Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSILLFRKKENKDKX 933

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4725	100.0	933	2 S41539	fibrinogen-binding
2	4399	93.1	989	2 D89852	fibrinogen-binding
3	1638.5	34.7	877	2 F90070	Clumping factor B
4	1628	34.5	1092	2 T30214	fibrinogen-binding
5	1414	29.9	1166	2 T28680	fibrinogen-binding
6	1395.5	29.5	1385	2 D89824	hypothetical prote
7	1354.5	28.7	953	2 C89824	hypothetical prote
8	1320.5	27.9	1141	2 E89824	hypothetical prote
9	1110.5	23.5	1315	2 T28679	fibrinogen-binding
10	963	20.4	903	2 AG1299	probable peptidogl
11	957.5	20.3	882	2 AG1671	probable peptidogl
12	799	16.9	2271	2 F90073	hypothetical prote
13	688.5	14.6	961	2 G90053	hypothetical prote
14	682	14.4	940	2 S19702	fibrinectin-bindin
15	659.5	14.0	2570	2 T17451	fimbriae-associate
16	635.5	13.4	406	2 S38170	SRP40 protein - ye
17	614	13.0	1018	2 A32192	fibrinectin-bindin
18	611	12.9	1038	2 H90053	hypothetical prote
19	597.5	12.6	640	2 A54502	S antigen precursor
20	577	12.2	4776	2 E95206	cell wall surface
21	569.5	12.1	1063	2 D86731	hypothetical prote
22	500	10.6	334	2 A54138	acidic repetitive
23	494	10.3	1428	2 T08850	lustrin A - Califo
24	487.5	10.3	3394	2 T18501	hypothetical prote
25	479.5	10.1	1664	2 T18262	S-layer protein -
26	461	9.8	1217	2 S52714	sericinB - silkw
27	445	9.4	888	2 T46726	secreted acid phos
28	430.5	9.1	1459	2 T32271	hypothetical prote
29	426.5	9.0	833	2 E90577	lipoprotein vsaC [

30	420	8.9	792	2 T42963	hypothetical prote
31	407	8.6	2232	2 T34434	hypothetical prote
32	405.5	8.6	937	2 S58135	hypally regulated
33	401	8.5	2910	2 T28156	DNA-directed RNA p
34	388	8.2	286	2 C61615	sericin MG-2 - gre
35	379	8.0	1072	2 A86827	hypothetical prote
36	378	8.0	797	2 A36811	hypothetical prote
37	377.5	8.0	2481	2 D90011	FmtB protein (impo
38	377	8.0	1192	2 A71623	probable secreted
39	374	7.9	1419	2 T30531	agglutinin-like ad
40	374	7.9	1829	2 T24563	hypothetical prote
41	373.5	7.9	670	2 T28391	ORF MSV230 hypothe
42	372.5	7.9	695	2 S27390	calcium-binding pr
43	368.5	7.8	955	2 T18435	hypothetical prote
44	363	7.7	489	2 A45988	dentin matrix acid
45	363	7.7	685	2 E71609	hypothetical prote

ALIGNMENTS

RESULT 1

S41539

fibrinogen-binding protein - Staphylococcus aureus

N;Alternate names: clumping factor

C;Species: Staphylococcus aureus

C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S41539, 836630

R;McDewitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus

A;Reference number: S41539; MUID: 94224142; PMID: 8170386

A;Accession: S41539

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-933 <MCD>

A;Cross-references: UNIPROT: Q53653; EMBL: Z18852; NID: g397525; PID: g39752

Query Match 100.0%; Score 4725; DB 2; Length 933;

Best Local Similarity 100.0%; Pred. No. 1.4e-202;

Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNMKKEKHAIRKKSIGVASVLGTLIGFGLLSKEADASENSVTQSDSASNEKSDSS	60
DB	1	MNMKKEKHAIRKKSIGVASVLGTLIGFGLLSKEADASENSVTQSDSASNEKSDSS	60
QY	61	SVSAAPKTDITVSDTKTSNTNNGETSVAQNPAAQETTQSSSTNATTEPTVTGEATTT	120
DB	61	SVSAAPKTDITVSDTKTSNTNNGETSVAQNPAAQETTQSSSTNATTEPTVTGEATTT	120
QY	121	TTNQANTPATQSSNTNABELVNQTSNETTFNDTNTVSSVNSPQNSNAENVSTQDTST	180
DB	121	TTNQANTPATQSSNTNABELVNQTSNETTFNDTNTVSSVNSPQNSNAENVSTQDTST	180
QY	181	EATPSNNEAPQSTDAASNDVNVQAVNTSAPRMEAFSLAAVAADAPAAAGTDITNQLTNVT	240
DB	181	EATPSNNEAPQSTDAASNDVNVQAVNTSAPRMEAFSLAAVAADAPAAAGTDITNQLTNVT	240
QY	241	VGIDSGTIVYVPHQAGVYKLVNGFSPVNSAVKGTDFKIVTPEKELNGLVSTAKVPPIMAG	300
DB	241	VGIDSGTIVYVPHQAGVYKLVNGFSPVNSAVKGTDFKIVTPEKELNGLVSTAKVPPIMAG	300
QY	301	DOVLANGVSDSGNVLYTFTDYNTKDDVKATLTMPAYIDPENVKKTGNVLTATGISTT	360
DB	301	DOVLANGVSDSGNVLYTFTDYNTKDDVKATLTMPAYIDPENVKKTGNVLTATGISTT	360
QY	361	ANKTVLVDVEKYGKFNLSIKGTDIDKTNNTYRQTIYVNPQSDNVIAPLVGLNKLKPT	420
DB	361	ANKTVLVDVEKYGKFNLSIKGTDIDKTNNTYRQTIYVNPQSDNVIAPLVGLNKLKPT	420
QY	421	DSNALIDQQNTSIKVKVDNAADLSSEYFVNPENFEDVTNSVNTITFPNPQYKVEFTPD	480
DB	421	DSNALIDQQNTSIKVKVDNAADLSSEYFVNPENFEDVTNSVNTITFPNPQYKVEFTPD	480

```
QY 481 DOITPTIIVVNGHIDPNSKGDALRLSTLYGNSNIIWRSMWNEVAFNNGSGSGDGID 540
Db 481 DOITPTIIVVNGHIDPNSKGDALRLSTLYGNSNIIWRSMWNEVAFNNGSGSGDGID 540
QY 541 KPVPEQDEPEGEIEPIPEDSDPGSDGSDNSDSDSGSDGSDSTSDSGSDSASDSDSA 600
Db 541 KPVPEQDEPEGEIEPIPEDSDPGSDGSDNSDSDSGSDGSDSTSDSGSDSASDSDSA 600
QY 601 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 660
Db 601 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 660
QY 661 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 720
Db 661 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 720
QY 721 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 780
Db 721 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 780
QY 781 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 840
Db 781 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 840
QY 841 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
Db 841 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
QY 901 SEDEANTSLIWGLLASIGSLILFRKKENKDKK 933
Db 901 SEDEANTSLIWGLLASIGSLILFRKKENKDKK 933

RESULT 2
D89852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D89852
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: D89852
A>Status: preliminary
A:Molecule type: DNA
A:Cross-references: UNIPROT:Q99WJ4; GB:BA000018; PID:g13700678; PIDN:BABA1975.1; GSPDB:G
A:Experimental source: strain N315
A:Genetics:
A:Gene: c1fa

Query Match 93.1%; Score 4399; DB 2; Length 989;
Best Local Similarity 88.6%; Pred. No. 4.6e-188;
Matches 876; Conservative 27; Mismatches 30; Indels 56; Gaps 2;
QY 1 MMKKKEKHAIKKIGVASVLVGLIGFLLSSKEADASENSVTQSDASNESKSDSS 60
Db 1 MMKKKEKHAIKKIGVASVLVGLIGFLLSSKEADASENSVTQSDASNESKSDSS 60
QY 61 SVSAAPKTDITNVSDTKSSNTNNGETSVAQPAQETTQSSSTNATTEETPVTEATT 120
Db 61 SVSAAPKTDITNVSDTKSSNTNNGETSVAQPAQETTQSSSTNATTEETPVTEATT 120
QY 121 TTQANTPATQSSNTNABELVNOTSETTFNDTNTVSVNSPQNSNTAENVSTTQDST 180
Db 121 TTQANTPATQSSNTNABELVNOTSETTFNDTNTVSVNSPQNSNTAENVSTTQDST 180
QY 181 EATPNNESAPQSDASNDKDVNOAVNTSAPRMRAFLAAVAADAPAAAGTDTITNQLTNYT 240
Db 181 EATPNNESAPQSDASNDKDVNOAVNTSAPRMRAFLAAVAADAPAAAGTDTITNQLTNYT 240
```

```
Db 181 EATPNNESAPQSDASNDKDVNOAVNTSAPRMRAFLAAVAADAPAAAGTDTITNQLTNYT 240
QY 241 VGIDSGTTVYPHQAGYVKLYGFSVPNSAVAGDTEFKIIVPKELNLNGVTSTAKVPPIIMAG 300
Db 241 VTIIDSGTTVYPHQAGYVKLYGFSVPNSAVAGDTEFKIIVPKELNLNGVTSTAKVPPIIMAG 300
QY 301 DQVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVLTATGISTT 360
Db 301 DQVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVLTATGISTT 360
QY 361 ANKTVLVYKFKFYNLSIKGTIDQIDKTNNTYQTIYVNPNGNVIAPVLITGNLKPT 420
Db 361 ASKTVLIDYKFGQPHNLSIKGTIDQIDKTNNTYQTIYVNPNGNVIAPVLITGNLKPT 420
QY 421 DSNALIDQNTSIIKVKYKVDNAADLSESYFVNPFENFEDVTNSVNIITFPNPNQYKVFENPTD 480
Db 421 KSNALIDAKNTDIKVRVDNANDLSESYFVNPFENFEDVTNQVRISFPNANQYKVFETDD 480
QY 481 DQITPTIIVVNGHIDPNSKGDALRLSTLYGNSNIIWRSMWNEVAFNNGSGSGDGID 540
Db 481 DQITPTIIVVNGHIDPASTGDLALRLSTFYGYDSNFIWRSMWNEVAFNNGSGSGDGID 540
QY 541 KPVPEQDEPEGEIEPIPEDSDPGSDGSDNSDSDSGSDGSDSTSDSGSDSASDSDSA 600
Db 541 KPVPEQDEPEGEIEPIPEDSDPGSDGSDNSDSDSGSDGSDSTSDSGSDSASDSDSA 600
QY 601 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 622
Db 601 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 660
QY 623 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 676
Db 661 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 720
QY 677 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 736
Db 721 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 780
QY 737 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 796
Db 781 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 840
QY 797 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 853
Db 841 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 900
QY 854 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 904
Db 901 STSDTGSNDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 960
QY 905 ANTSLIWGLLASIGSLILFRKKENKDKK 933
Db 961 ANTSLIWGLLASIGSLILFRKKENKDKK 989
```

```
RESULT 3
F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: F90070
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <KUR>
A:Cross-references: UNIPROT:Q99R07; GB:BA000018; PID:g13702588; PIDN:BABA3728.1; GSPDB:G
A:Experimental source: strain N315
```


C;Genetics:
A;Gene: clfb

```
Query Match 34.7%; Score 1638.5; DB 2; Length 877;
Best Local Similarity 40.4%; Pred. No. 1.3e-65;
Matches 398; Conservative 150; Mismatches 262; Indels 175; Gaps 33;

QY 5 KKEKHAIRKKSIGVASVLGTLIGFGLLSKEADASE--NSVTQSDASNESKNSDSSV 62
DB 11 KQKYSIRRTVTGTSVIGATILFG-IGNHQQAQSEQSDNTTQS-SKNWASADSEKNM 68
QY 63 SAAPKTDITNVSDTKSSNTNNGETSAQNPAAQOETTQSSSTN-ATTEPTPTVGEATTTT 121
DB 69 IETPQLNTTANDSDISANTNSANVDSTTKPMSTQTSNTTTEPASNETP-----QPTAI 124
QY 122 TNOANTPATQSSNTNAELVNQTSNETTENDTNTVSSVNSPQNSNAENVSTQDTSTE 181
DB 125 KNOAT--AAKMQDQTVFQEAANSQVDNKT--NDANSI-----ATNSE-----LKNSQTL 169
QY 182 ATPSNNEAPQSDASNDKDVNOAVNTSAPRMRAFSLAAVAADAP-----AAGTDTNQ 235
DB 170 DLP--QSSPQT-----ISNAQTSKPSVTRAVRSLAVAEPPVNAADAKGTNVNDK 218
QY 236 LTNVTVIGDSGTTVYHQAGYVKLNYGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVP 295
DB 219 VTASNFKLEK-TTFDPNQSGNTFMAANFTVTDVKYSGDYF-----TAKLP 262
QY 296 PIMAGQVLANGVID-----SDGNVI-----YTFDTYVNTKD 327
DB 263 -----DSLNGNGVDVYSSNNTMPIADIKSTNGDVAKATYDILTKTFTVFVDYVNKE 317
QY 328 DVKATITMPAYIDPENVKKTGNVTLATIGSTTANKTVLDYDEK---YKGFYNLSIKGT 383
DB 318 NINGQSLPFTBRAPKPSGTYDANINADIENFNKITYNSPIAGIDKPNGANISSQ 377
QY 384 IDQIDKTN--NTYKQTIYVNS-----GDNVIAVPLTGNLKNPTDSNALIDQNTSIKVYK 437
DB 378 IIGVDTSAGQNTYKQTVFVNPQKQVLGNTWV--YIKGYQDKIESSGKVSATDKLRIPE 435
QY 438 VDMAADLSESYFNP--ENFEDVTNSV--NITFPNPQYKVENPTPDQITTYIIVVNG 493
DB 436 VNTSKLSLSDYADPNDSNLKEVTDQPKRIYYEHPNVAISKFG-----DITKYVWLVEG 491
QY 494 HIDPNKSGDLALSTLYGN-----SNTIWRSMGMDNEVAFNNGSGSGDGDKPVPBPQPD 549
DB 492 HYDNTGKN---LKTQVLOENVDPVTRNDYSIFGNWNNVYVYGGGSADG--DSAVNPXDP- 546
QY 550 EPGIEPIPEDSDSDPGSDSGSDNSDSDSGSDSGSDSTSDSGSDSASDSDSASDS 609
DB 547 TPG--PPVDPEPDP-----EPEPTPDP 568
QY 610 DSASDSDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 669
DB 569 EPGPDEPEPEPDPDPDPDSDSGSDSGSDSGSDSGSDSGSDSDSDSDSDSDSDSDSDS 628
QY 670 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 729
DB 629 ESUSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 688
QY 730 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 789
DB 689 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 748
QY 790 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 849
DB 749 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 795
QY 850 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 908
DB 796 -SDSDSDSDSR---VTPPNNEQKAPSNPKGEVNHNSKVSKQHTDALPETGDKSENTNAT 851
QY 909 LIWGLLASIGSLILFRKKENKOKK 933
```

DB 852 LFGAMMALLGSLILFRKKQDHKEK 876

RESULT 4
T30214

fibrinogen-binding protein - Staphylococcus epidermidis
C;Species: Staphylococcus epidermidis

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C;Accession: T30214

R;Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998

A;Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.

A;Reference number: Z20781; MUID:98261511; PMID:9596732

A;Accession: T30214

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1092 <NIL>

A;Cross-references: UNIPROT:O70022; EMBL:Y17116; NID:el296734; PID:el296735; PIDN:CAA7663

Query Match 34.5%; Score 1628; DB 2; Length 1092;

Best Local Similarity 36.8%; Pred. No. 5e-65;

Matches 421; Conservative 141; Mismatches 297; Indels 284; Gaps 32;

QY 5 KKEKHAIRKKSIGVASVLGTLIGFGLLSKEADASENSV-----TQSDASN 52

DB 18 KSNKYAIRKFTVGTASIVIGATLLFG-LGHNEAKAEENSVDVKDSNTDDELSDSDQSS 76

QY 53 ESKSNSSSVSAAPKTDITNV-----SDTKTSNTN---NGETSVAQNP 93

DB 77 DEEKNNVINNQISNTDDNNQIIKKEETNNYDGIKRSDETRTESTTNVDENEATFLQKTP 136

QY 94 -----AQOETQSSSTNATTEPTVGTGATTTTNOANTPATQOS-SNTNAELVN--- 143

DB 137 QDNTHLTEEVEKSSSVSSNSSSIDTAQSPSHITINREESVQSDNVEDSHVDFANSKI 196

QY 144 QTSNETTFNDNTV-----SSVNSPQNSNAENVSTQDTSTEATPSN---NESAP 191

DB 197 KESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELLN-LPINEVENKARP 255

QY 192 QSTDASKOVNOAVNTSAPRMRAFSLAAVAADAPAGTDTITNLTNTVTGI-DSGTTVY 250

DB 256 LSTTSAGPSIKRYTVN-----QLAAEQGSNVNHLIKVTDQ---SITEGYDDSEGVK 304

QY 251 PHQAGVVKLNYGFSVPNSAVKGFKITVPKELNGLNGVTSTAKVPPIMAGD-OVLANGVI 309

DB 305 AHDAENLIYDVTFEVDKVKSGDTMTVDIDKNVPSDLTDSFTPIPKIKNSGRIIATGY 364

QY 310 DS-DGNVIYFTFDYVNTKDDVKATLTMPAYIDPENVKKTG---NVTIATGIGSTTANKTV 365

DB 365 DNKNKQITVYFTDYVKYENIKAHKLKITSYDKSKVPNNNTKLDVEYKTALSS--VNKTI 422

QY 366 LVQYKYGKFEYNLSIKGTIDQDKTNTYQTIYVNPNGDNVIAVPLTGNLKNPTDSNAL 425

DB 423 TVEYQRPNERNTANLQSMFTNIDTKHVTVEQTIYINPL--RYSAKETNWNISNGDEGST 480

QY 426 IDQOQNTSIVKYVDNAADLSESYFV-NPENFEDVTNSVNITFPNPQYKVEFNTPDQIT 484

DB 481 IIDDSIIIKYKVDNQNLDPNSRIYDYSEYEDVTDNDYLAQLGNNDVNINFG-----NID 536

QY 485 TPYIVVNGHIDPNSKGD-----LALRSTLYGYNSNIIWRSMGMDNEVAFNNGSGSD 537

DB 537 SPVILKISKYDEN-KDDYTTIQQVTMTQTTINEYGE--FRTASYDNTAFSTSSGQG 593

QY 538 G-----IDKPVVP-----EQP-----DEGEIE 555

DB 594 GDLPEPKTYKIGDYVWEDVDKGIQNTNDNEKLSNVLVTLTYPDGTSKVRDEDEGKYQ 653

QY 556 -----PIPEDSDSDPGSDSGSDS-----NSDSGSDG----- 582

DB 654 FDGLKNGLYTKITFTPEGYTPTLKHSNTPALDSEGNVWVVTINGODMTIDSGFYQTP 713

QY 583 ----- 582

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: D89824

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1385 <KUR>

A;Cross-references: UNIPROT:Q99W47; GB:BA000018; PID:gl3700454; PIDN:BA041751.1; GSPDB:G

A;Experimental source: strain N315

C;Genetics:

A;Gene: sdrC

Query Match 29.5%; Score 1395.5; DB 2; Length 1385;

Best Local Similarity 29.6%; Pred. No. 1.4e-54;

Matches 416; Conservative 149; Mismatches 329; Indels 513; Gaps 40;

```
QY 1 MNMKKEKHAIRKKSIGVASVLVGTLLIGFGL--LSSKEADASENSVTQSDASNESKSN 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 15 MVSRLNKFIRKYVTGVTASILVGTLLIFGLNQEAKEASTNKELNEATTSASDNQSSD 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 SSSVSAAPKDDT--NVSDTKTSNTNNGETSVAGNPAQOETQSSNTNATEETPTVGEA 117
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 75 KVMQQLNQEDNTKNDKQKMWSSQGNERTSNGNKSIKESVQSTGKNV-----EV 126
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 TTTTNCANPATTQSSNTNAEELVNQSTNETTFDNTVSSV---NSPQNSTNAENVST 174
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 STAKSDQCAKPKSTNEDLNTKQIISNQEGLQPDLLNKSVNVQPTNEENKKVDKATTEST 186
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 TQDTSTBATPNNESAPOSTDAGN-----KOVVQAQVNTSAPRMAFSLAAVAAD 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 TLNVKSDAKSNAETLVDDNNSNNENNADIILPKSTAPKSLNT---RMRMAAIQPNSTD 243
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 APAAGTDITNQLNVTGIDSGITVYVPHQAGYKVLNFGFVSPNSAVKGDFTFKITVPKELN 284
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 SKNVNDLITNTLTVVVDADNSKTIVPAQ--DYLKSKSQITVDDKVKSGDYFTIKYSPTVQ 302
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 LMGVT-----STAKVPPIMAGDQVLANGVIDSDGNI--YTFDYVNTKDDVKATLTMP 336
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 303 VYGLNPEDIKNIQDIKDP---NNGETATAXHDTANNLLIYTFDYVDRFNSVKGNGINS 359
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 337 AYIDPENV---KKTGNVLTATGISTTANKTVLVDEYKYGKFNLSIKG---TIDQIDK 389
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 IYMDADTIPVDKK--DVPFSTVINGQITTTTADITYPAYKEADNNSIGSAFTETVSHVN 417
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 390 TNNT--YRQIYVNPSSGNI--APVLGNLKPNTDSN-ALIDQNTTSIKYKVDNADLS 445
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 418 VEDPGYNQVYVNPMDKDLGAKLVEAYHPKYPTNIGQINQVNTIKIYRVEGYTLN 477
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 446 ESYFVNPENFEDVTN-----SVNITP----- 466
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 478 KGYDVNTDLVDVTDEPKNMYKNSQSNVLDLFGDITSAVVMYNTKFOYNSESPTLVQ 537
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 ----- 466
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 538 MATLSSTGNKSVSTGNALGFTNNSQSGAGQEVYKIGNYVWEDTNKNGVQELGKGVNVT 597
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 -----PN--PN--QYKVEF--NTPDDQITTP----- 486
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 598 VTFVDDNNTNKVGEAVTKEDGSYLIPNLPGDYRVFESNLPKGYEFTPSKQGNNEELDSN 657
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 487 ---YIVVVGHDIPNSGDILALRSTLVGYNSNIWRSMWDEVAFNNKSGSGD--GIDK 541
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 658 GLSSVITVNG--KDNLSADLGIYKPKYKL--GDYVWEDTN-----KNGIQDQDEKIGS 707
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 542 FVYVEQDEPGEIEPIEDSDS-----DPGSDSGSDSN 574
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 708 VTVTLKDNQVNLKVTITDADGKYKFTDLDNQYKVEFTTPEGYTPPTVTSKSDIEKDSN 767
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 575 -----SDSGSDSGSDS----- 587
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 768 GLITTTGVINGADNMTLDSGFKYTPKYNLGNVWEDTNKQDQSTKTEKIGSVTVTLKNEN 827
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 588 -----DSGSDSASDSASDSASDSAS 607
Db 828 GEVLQTTKTDCKGKYQFTGLENGTYKVEFTPSGYTPTQVSGTDEGIDNGTSTTGVIK 887
QY 608 DSDS-----ASDSASDS----- 622
Db 888 DKNDNTIDSGFYKFTYNLGDYVWEDTNKNGVQDKDEKIGSVTVTLKXENDKVLKTVTTD 947
QY 623 -----SNDSDSDSD-----SDSDS----- 637
Db 948 ENKGYQFTDLNNGTYKVEFTPSGYTPTSVTSGNDTEKDSNGLTTTGVIKADNMTLDSG 1007
QY 638 -----DSDSDSDSDS----- 647
Db 1008 FYKTPKYSGLDYVWYDSNKGQKQDSTKGIKDYKVVILLNEKGEVIGTTKTDENGKYRFDN 1067
QY 648 -----DSDSDSD-----SDSD-----SDSDSDSD 666
Db 1068 LDSGKYKVIPEKFTGLTGTNTTETDDKADGGEVDVTITDHDFTLLDNGYEEETSDSD 1127
QY 667 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 726
Db 1128 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1187
QY 727 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 786
Db 1188 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1247
QY 787 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 846
Db 1248 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1307
QY 847 SDSSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 905
Db 1308 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1358
QY 906 -NTSLWGLLASIGSLILF--RRKKENK 930
Db 1359 NNATLFGGLFAALGSLILFGRKKQNK 1385

RESULT 7
C89824
hypothetical protein sdrC [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: C89824
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: C89824
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-953 <KUR>
A;Cross-references: UNIPROT:Q99W48; GB:BA000018; PID:gl3700453; PIDN:BA041750.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: sdrC

Query Match 28.7%; Score 1354.5; DB 2; Length 953;
Best Local Similarity 36.4%; Pred. No. 5.8e-53;
Matches 388; Conservative 132; Mismatches 298; Indels 247; Gaps 39;

QY 1 MNMKKK-----EKHAIRKKSIGVASVLVGTLLIGFGLLSKEADASENSVTQSD 48
Db 1 MNMKKTATNRKGMIPNRLNKFIRKYSVGTASILVGTLLIFG--LSGHEAKAAEHTNGELN 59
QY 49 SAKNESKNSDSSVSAAPKTTDDTNVSDTKTSNTNNGETSVAGNPAQOETQSSNTNATT 108
Db 60 QSKNETTA--PSENKITEKVDRLQKD-----NTQTATADQP---KVTMSDS--ATV 104
```


C;Accession: AG1299
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1299
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-903 <GLA>
A;Cross-references: UNIPROT:Q8Y697; GB:NC_003210; PIDN:CAC99877.1; PID:g16411253; GSPD:
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1799

[illegible]

```

RESULT 11
AG1671
probable peptidoglycan bound protein (LPXTG motif) lin1913 [imported] - Listeria innocua
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG1671

```

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluerer, T.; Simoes, N.; Tierrez, A.; Vazquez-Roland, J.A.; Voss, H.; Wehlund,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AGL671
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-882 <GLA>
A:Cross-references: UNIPROT:Q92AK9; GB:AL592022; PIDN:CAC97143.1; PID:gl6414414; GSPDB:G
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1913

[illegible]

RESULT 12
F90073
hypothetical protein SA2447 [imported] - *Staphylococcus aureus* (strain N315)
C/Species: *Staphylococcus aureus*
C/Date: 10-May-2001 #sequence revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: F90073
R/R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, Y.; Ma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekikawa, C.; Hattori, M.; Hayashi, H.; Hiramatsu, K.
C/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
Lancet 357, 1225-1249, 2001

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 04:52:14 ; Search time 1562 Seconds

(without alignments)
11755.746 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaccataaattacacatc.....gaagagtataaagaaagctt 3498

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 23Sep04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3498	100.0	3498	5	Aaf58593 Staphyloc
2	3487	99.7	3499	3	Aaz55832 Staphyloc
3	3279.6	93.8	4709	2	Aav74898 Staphyloc
4	2797.4	80.0	2799	8	Abt14896 Pathogen
5	2746.6	78.5	2781	8	AcF73781 Staphyloc
6	2331.6	66.7	2808	8	ACA20272 Prokaryot
7	2096.8	59.9	3066	4	AAS54809 Staphyloc
8	2093.8	59.9	3063	4	AAS51834 Staphyloc
9	1776.4	50.8	2792	6	Abk13616 S. aureus
10	1776.4	50.8	2792	8	ADA89855 Staphyloc
11	1560	44.6	1560	6	Aad46861 Staphyloc
12	990	28.3	990	6	Aad46862 Staphyloc
13	619.2	17.7	5406	3	Aaz93533 Cell wall
14	619.2	17.7	5406	3	Aaz51201 Staph. ep
15	589.6	16.9	4627	4	Aah54581 S. epider
16	589.2	16.8	4899	8	ACA46383 Prokaryot
17	584.8	16.7	1485	4	Aah52321 S. epider
18	568.6	16.3	2369	2	Aax77591 S. aureus
19	568	16.2	2739	8	Abt14866 Pathogen
20	558.2	16.0	1059	6	AbS52617 Staphyloc
21	532.4	15.2	1539	8	ACA35100 Prokaryot

22	525.4	15.0	2151	8	ABT14861 Pathogen
23	510	14.6	2631	8	ACF74262 Staphyloc
24	510	14.6	3007	8	ADA89797 Staphyloc
25	506.4	14.5	2010	8	ACA46390 Prokaryot
26	502.6	14.4	2582	8	ADA89798 Staphyloc
27	502.6	14.4	2582	8	ADA89868 Staphyloc
28	502	14.4	2634	8	ACA46374 Prokaryot
29	497.8	14.2	11050	2	AAV74407 Staphyloc
30	485.6	13.9	4911	8	ABT15076 Pathogen
31	462	13.2	3600	2	AAV04279 Staphyloc
32	448.4	12.8	2985	8	ACF73997 Staphyloc
33	446	12.8	2428	8	ADA89818 Staphyloc
34	439.4	12.6	4158	8	ACA20270 Prokaryot
35	436.8	12.5	756	4	AAH52319 S. epider
36	436.6	12.5	4143	8	ABT14918 Pathogen
37	415	11.9	5764	8	ACA47359 Prokaryot
38	383.6	11.0	4418	2	AAV74765 Staphyloc
39	375.6	10.7	2796	4	AAS51941 Staphyloc
40	375.6	10.7	2799	4	AAS54703 Staphyloc
41	365	10.4	396	2	AAV77913 Staphyloc
42	361.4	10.3	2841	8	ABT14889 Pathogen
43	360.6	10.3	3498	8	ABT14929 Pathogen
44	360.6	10.3	4005	2	AAV74942 Staphyloc
45	360.6	10.3	4047	8	ACF73996 Staphyloc

ALIGNMENTS

RESULT 1

AAF58593

ID AAF58593 standard; DNA; 3498 BP.

AC AAF58593;

DT 23-APR-2001 (first entry)

DE Staphylococcus aureus clfa gene.

KW Staphylococcus aureus; clfa; antibiotic; vaccine;
KW fibrinogen binding protein; bacterial infection; mastitis; ds.

OS Staphylococcus aureus.

PN US6177084-B1.

PD 23-JAN-2001.

PF 19-OCT-1999; 99US-00421868.

PR 22-AUG-1994; 94US-00293728.

XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

XX Foster TJ, Mcdevitt DL;

XX WPI: 2001-181608/18.

XX P-PSDB; AAB69508.

XX Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
XX to prevent infection, promote wound healing, block adherence to
XX indwelling medical devices and for diagnosing staphylococcus aureus
XX infection.

XX Claim 1; Fig 2; 30pp; English.

XX The present sequence encodes a novel Staphylococcus aureus fibrinogen
XX binding protein. It is useful as a vaccine to protect against human and
XX animal infections caused by S. aureus, such as against mastitis, to block
XX S. aureus from colonising and infecting a wound, to block adherence of
XX S. aureus to indwelling medical devices such as catheters, replacement
XX heart valves and cardiac assist devices. The protein can be used to
XX diagnose bacterial infections. Polyclonal and monoclonal antibodies

Oppi, C.R.

CC raised against the fibrinogen binding protein are useful for passively
CC immunising against infections caused by S aureus, to prevent infection of
CC a wound and to diagnose bacterial infections

XX	SQ	Sequence 3498 BP; 1260 A; 587 C; 704 G; 947 T; 0 U; 0 Other;	
		Query Match	
		Best Local Similarity 100.0%; Score 3498; DB 5; Length 3498;	
		Matches 3498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	GGTACCAATAATACACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATTAACATTA	60
Db	1	GGTACCATAAATACACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATTAACATTA	60
Qy	61	GGTAGAGTTTCATATTAATAATAAATAATGTTTGAACAAATCGTAGCTTGTCTGTTGTA	120
Db	61	GGTAGAGTTTCATATTAATAATAAATAATGTTTGAACAAATCGTAGCTTGTCTGTTGTA	120
Qy	121	ATTCTTAAATAGCAATTAATAATAAATGTTTGTAGTAAAGTATTAATTTGGATAATAAAA	180
Db	121	ATTCTTAAATAGCAATTAATAATAAATGTTTGTAGTAAAGTATTAATTTGGATAATAAAA	180
Qy	181	TATCGATACAAATTAATGCTATTAATGCAATTTTAGTGTATTAATTCATTAACAGAGATT	240
Db	181	TATCGATACAAATTAATGCTATTAATGCAATTTTAGTGTATTAATTCATTAACAGAGATT	240
Qy	241	AAATATATCTTAAAGGCTATATAGTTAATATAAATGACTTTTTAAAGAGGGAATAAA	300
Db	241	AAATATATCTTAAAGGCTATATAGTTAATATAAATGACTTTTTAAAGAGGGAATAAA	300
Qy	301	ATGAATATGAAGAAAAAGAAAAACAGCAATTCGGAATAAATCGATTGGCGTGCCTCA	360
Db	301	ATGAATATGAAGAAAAAGAAAAACAGCAATTCGGAATAAATCGATTGGCGTGCCTCA	360
Qy	361	GTGCTTAGTAGTACGTTAAATCGGTTTTGGACTACTACGAGTAAAGACAGATGCAAGT	420
Db	361	GTGCTTAGTAGTACGTTAAATCGGTTTTGGACTACTACGAGTAAAGACAGATGCAAGT	420
Qy	421	GAAATAGTCTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAGAT	480
Db	421	GAAATAGTCTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAGAT	480
Qy	481	AGCGTTAGTCTGCACCTAAACAGACACAAACGCTGAGTGATCTAAACACATGCTCA	540
Db	481	AGCGTTAGTCTGCACCTAAACAGACACAAACGCTGAGTGATCTAAACACATGCTCA	540
Qy	541	AAACATTAATGCGGAAACGAGTGTGCGCAAAATCCAGCACAACAGGAAACGACAAA	600
Db	541	AAACATTAATGCGGAAACGAGTGTGCGCAAAATCCAGCACAACAGGAAACGACAAA	600
Qy	601	TCATCATCAAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTACG	660
Db	601	TCATCATCAAAATGCAACTACGGAAGAAACGCGGTAACCTGTTGAAGCTACTACTACG	660
Qy	661	ACAAAGAAATCAAGTAAATACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA	720
Db	661	ACAAAGAAATCAAGTAAATACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA	720
Qy	721	TTAGTGAATCAAAACAGTAATGAACGACTTTTAATGATTAATACAGTATCATCTGTA	780
Db	721	TTAGTGAATCAAAACAGTAATGAACGACTTTTAATGATTAATACAGTATCATCTGTA	780
Qy	781	AAITTCACCTCAAAATCTCAAAATGCGAAATGTTTCAACAAACGCAAGTACTTCAACT	840
Db	781	AAITTCACCTCAAAATCTCAAAATGCGAAATGTTTCAACAAACGCAAGTACTTCAACT	840
Qy	841	GAAGCAACACCTTCAACAAATGATCGTCCACAGAGTACAGTGAAGTAAATAAGAT	900
Db	841	GAAGCAACACCTTCAACAAATGATCGTCCACAGAGTACAGTGAAGTAAATAAGAT	900
Qy	901	GTAGTTAATCAACGGGTTAATACAAAGTGCCTTAGATGAGCATTTAGTTTAGCGGCA	960
Db	901	GTAGTTAATCAACGGGTTAATACAAAGTGCCTTAGATGAGCATTTAGTTTAGCGGCA	960

Qy	961	GTAGTCAGATGCACCGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGACA	1020
Db	961	GTAGTCAGATGCACCGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGACA	1020
Qy	1021	GTTGGTATTGACTCTGCTAGCACTGTGTATCCGCCAACAGCAGGTTATGTCAAACTGAAT	1080
Db	1021	GTTGGTATTGACTCTGCTAGCACTGTGTATCCGCCAACAGCAGGTTATGTCAAACTGAAT	1080
Qy	1081	TATGGTTTTTTCAGTGCCTAAATTTCTGCTGTTTAAAGGTGACACATTCAAAATACCTGACCT	1140
Db	1081	TATGGTTTTTTCAGTGCCTAAATTTCTGCTGTTTAAAGGTGACACATTCAAAATACCTGACCT	1140
Qy	1141	AAAGAAATTAACCTTAAATGGTAACTTTCAACTGCTAAAGTGCACCAATTTATGGCTGGA	1200
Db	1141	AAAGAAATTAACCTTAAATGGTAACTTTCAACTGCTAAAGTGCACCAATTTATGGCTGGA	1200
Qy	1201	GATCAAGTATTGGCAATGGTAAATCGATAGTAGTAAATGTTTATTAACATTTTACA	1260
Db	1201	GATCAAGTATTGGCAATGGTAAATCGATAGTAGTAAATGTTTATTAACATTTTACA	1260
Qy	1261	GACTATGTAATTAATAAGATGATGTAAGAGCACTTTGACCATGCCGCTTATATTGAC	1320
Db	1261	GACTATGTAATTAATAAGATGATGTAAGAGCACTTTGACCATGCCGCTTATATTGAC	1320
Qy	1321	CCTGAAATCTTAAAGAGACAGGTAATGTGACATTTGCTACTTGSCATAGGTAGTACAACA	1380
Db	1321	CCTGAAATCTTAAAGAGACAGGTAATGTGACATTTGCTACTTGSCATAGGTAGTACAACA	1380
Qy	1381	GCAAAACAAAACAGTATTAGTAAATATGTAAGTTTATTAACCTTATCTATT	1440
Db	1381	GCAAAACAAAACAGTATTAGTAAATATGTAAGTTTATTAACCTTATCTATT	1440
Qy	1441	AAAGGTACAAATTCACCAATTCGATTAACAAATAATACGTATCGTCAGACAATTTATGTC	1500
Db	1441	AAAGGTACAAATTCACCAATTCGATTAACAAATAATACGTATCGTCAGACAATTTATGTC	1500
Qy	1501	AAATCCAAAGTGCAGATACGTTATTGGCCGGTTTAAACAGGTAATTTAAACCAAAATACG	1560
Db	1501	AAATCCAAAGTGCAGATACGTTATTGGCCGGTTTAAACAGGTAATTTAAACCAAAATACG	1560
Qy	1561	GATAGTAATGCAATTAATAGATCAGCAAAATACAAATTAAGTATATAAAGTAGATAAT	1620
Db	1561	GATAGTAATGCAATTAATAGATCAGCAAAATACAAATTAAGTATATAAAGTAGATAAT	1620
Qy	1621	GCAGCTGATTTATCTGAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACATAAT	1680
Db	1621	GCAGCTGATTTATCTGAAGTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCACATAAT	1680
Qy	1681	AGTGTGAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATACGCTGAT	1740
Db	1681	AGTGTGAATATTACATTTCCCAATCCAAATCAATATAAAGTAGAGTTTAAATACGCTGAT	1740
Qy	1741	GATCAAAATTAACAACCGTATATAGTAGTTGTTAATAGTGCATATTTGATCCGAATAGCAAA	1800
Db	1741	GATCAAAATTAACAACCGTATATAGTAGTTGTTAATAGTGCATATTTGATCCGAATAGCAAA	1800
Qy	1801	GGTGATTTAGCTTTAGTTCAACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT	1860
Db	1801	GGTGATTTAGCTTTAGTTCAACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT	1860
Qy	1861	ATGTCATGGGACACCAAGTAGCATTTAATAACGGATCAGGTTCTGCTGACCGTATCGAT	1920
Db	1861	ATGTCATGGGACACCAAGTAGCATTTAATAACGGATCAGGTTCTGCTGACCGTATCGAT	1920
Qy	1921	AAACCAAGTTGTTCTGAAACCAACCTGATGAGCTGGTGAATTTGAAATTCAGAGGAT	1980
Db	1921	AAACCAAGTTGTTCTGAAACCAACCTGATGAGCTGGTGAATTTGAAATTCAGAGGAT	1980
Qy	1981	TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGCAGGATTTCTAATTCAGATACGGTTTCAGAT	2040
Db	1981	TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGCAGGATTTCTAATTCAGATACGGTTTCAGAT	2040

XX This sequence represents the *Staphylococcus aureus* *clfA* gene which
CC encodes a fibrinogen-binding protein, ClfA. ClfA is an important receptor
CC involved in *S. aureus* colonisation of indwelling medical devices (e.g.,
CC catheters, artificial heart valves). Shortly after implantation, the
CC surfaces of medical devices become coated with host plasma and matrix
CC proteins such as fibrinogen and fibronectin, and there is considerable
CC evidence to suggest that bacterial adherence to fibrinogen/fibrin is
CC important in the initiation of device-related infection. The fibrinogen-
CC binding region of ClfA is thought to reside between residues 332 and 550
CC in a region designated A. The protein also contains a repeated region
CC (region R) comprising 154 repeats of the dipeptide Ser-Asp, and the C-
CC terminus contains features present in surface proteins of other Gram
CC positive bacteria that are responsible for anchoring the protein to the
CC cell wall and cell membrane. ClfA, or its fragments, may be used to block
CC *S. aureus* colonisation of wounds, to prevent adherence of *S. aureus* to
CC indwelling medical devices, as vaccines to protect against *S. aureus*
CC infection (e.g., mastitis in ruminants), to raise specific antibodies,
CC and for diagnosis (by agglutination or immunoassay). The specific
CC antibodies are used for passive immunisation, to block infection of
CC wounds or adhesion of *S. aureus* and for diagnosis. Nucleotides encoding
CC ClfA and its fragments may be used as diagnostic probes
XX

SQ Sequence 3499 BP; 1260 A; 587 C; 704 G; 948 T; 0 U; 0 Other;
Query Match 99.7%; Score 3487; DB 3; Length 3499;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3498; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGTACCAATAATACACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATACATTA 60
DB 1 GGTACCAATAATACACATCTGCTTTTGAAGAAATATGATTTCAAGTAGGATACATTA 60

QY 61 GGTAGCTTCATATTATATAAATAATGTTTGAAGTAAAGTATATTGTGGATAATAAAA 120
DB 61 GGTAGCTTCATATTATATAAATAATGTTTGAAGTAAAGTATATTGTGGATAATAAAA 120

QY 121 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 180
DB 121 ATTCTTAAATAGCAATAAATAAATGTTTGTAGTAAAGTATATTGTGGATAATAAAA 180

QY 181 TATCGATACAAATTAATGCTATAATGCAATTTTGTAGTAAAGTATATTGTGGATAATAA 240
DB 181 TATCGATACAAATTAATGCTATAATGCAATTTTGTAGTAAAGTATATTGTGGATAATAA 240

QY 241 AAATATATC-TTAAAGGATATAGTTAATATAAATGACTTTTAAAGACGGGATAA 299
DB 241 AAATATATCTTTAAAGGATATAGTTAATATAAATGACTTTTAAAGACGGGATAA 300

QY 300 AATGAATATCAAGAAAAAGAAAAACACGCAATTCGGAAAAATTCGATTCGGCGTTC 359
DB 301 AATGAATATCAAGAAAAAGAAAAACACGCAATTCGGAAAAATTCGATTCGGCGTTC 360

QY 360 AGTGCCTGTAGTACGTTAATCGCTTTTGGACTACTCAGCAGTAAAGACAGATGCAAG 419
DB 361 AGTGCCTGTAGTACGTTAATCGCTTTTGGACTACTCAGCAGTAAAGACAGATGCAAG 420

QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAGTAAAGTAAAG 479
DB 421 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAAGCAAGCAAGTAAAGTAAAG 480

QY 480 TAGCGTTAGTGTGCACTTAAACACAGACGACAAACGTCGAGTGATCTAAACATCGTC 539
DB 481 TAGCGTTAGTGTGCACTTAAACACAGACGACAAACGTCGAGTGATCTAAACATCGTC 540

QY 540 AATCACTAATAATGGCGAAACAGTGTGGCGAAATCCAGACACACAGGAACGACACA 599
DB 541 AATCACTAATAATGGCGAAACAGTGTGGCGAAATCCAGACACACAGGAACGACACA 600

QY 600 ATCATCATCAACAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 659
DB 601 ATCATCATCAACAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 660

QY 660 GACAAACGAATCAAGCTAATAACACCGCAACAACTCAATCAAGCAATACAAAATGGGAGGA 719
DB 661 GACAAACGAATCAAGCTAATAACACCGCAACAACTCAATCAAGCAATACAAAATGGGAGGA 720

QY 720 ATTAGTGAATCAAAACAAAGTAAATGAACGACTTTTAAATGATACATAATACAGTATCATCTGT 779
DB 721 ATTAGTGAATCAAAACAAAGTAAATGAACGACTTTTAAATGATACATAATACAGTATCATCTGT 780

QY 780 AAATTCACCTCAAAATTTTCAAAATGCGGAAATGTTTCAACACGCAAGTATCTCAAC 839
DB 781 AAATTCACCTCAAAATTTTCAAAATGCGGAAATGTTTCAACACGCAAGTATCTCAAC 840

QY 840 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 899
DB 841 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAAATAAGA 900

QY 900 TGTAGTTAATCAAGCGGTTTAAATCAAGTGCCTTAGAATGAGAGCATTTAGTTAGCGGC 959
DB 901 TGTAGTTAATCAAGCGGTTTAAATCAAGTGCCTTAGAATGAGAGCATTTAGTTAGCGGC 960

QY 960 AGTAGCTGCAGATGCACCGGCAGCTGSCACAGATATTACGAATCAGTTGACGATGTGAC 1019
DB 961 AGTAGCTGCAGATGCACCGGCAGCTGSCACAGATATTACGAATCAGTTGACGATGTGAC 1020

QY 1020 AGTTGGTATTGACTCTCGGTACGACTGTGTATCCGCACCAAGCAGGTTTATGTCAAACTGAA 1079
DB 1021 AGTTGGTATTGACTCTCGGTACGACTGTGTATCCGCACCAAGCAGGTTTATGTCAAACTGAA 1080

QY 1080 TTATGGTTTTTCAAGTGCCTAAATTTCTGCTGTAAAGGTGACACATTTCAAAATACTGAC 1139
DB 1081 TTATGGTTTTTCAAGTGCCTAAATTTCTGCTGTAAAGGTGACACATTTCAAAATACTGAC 1140

QY 1140 TAAAGAAATTAACCTTAAATCGTGTAACTTCAACTGCTTAAAGTGCACCAATTTAGGCTGG 1199
DB 1141 TAAAGAAATTAACCTTAAATCGTGTAACTTCAACTGCTTAAAGTGCACCAATTTAGGCTGG 1200

QY 1200 AGATCAAGTATTGGCAAAATCGTGTAACTCGATAGTGTAAAGTATTTTATACATTTAC 1259
DB 1201 AGATCAAGTATTGGCAAAATCGTGTAACTCGATAGTGTAAAGTATTTTATACATTTAC 1260

QY 1260 AGACTATGTAATACTAAAGATGATGTAAGCAAACTTTGACCAATGCCCGCTTATATGA 1319
DB 1261 AGACTATGTAATACTAAAGATGATGTAAGCAAACTTTGACCAATGCCCGCTTATATGA 1320

QY 1320 CCTGGAATAATGTTAAAGACAGTAAATGTGACATTTGGCTACTGGCATAGTAGTACAAAC 1379
DB 1321 CCTGGAATAATGTTAAAGACAGTAAATGTGACATTTGGCTACTGGCATAGTAGTACAAAC 1380

QY 1380 AGCAAAACAAACAGTATTTAGTAGATTATGAAAAATATGTTAAAGTATTTTAACTTATCTAT 1439
DB 1381 AGCAAAACAAACAGTATTTAGTAGATTATGAAAAATATGTTAAAGTATTTTAACTTATCTAT 1440

QY 1440 TAAAGGTACAAATTTGACCAAAATCGATAAAACAAATAATACGTATCGTCAGACAATTTATGT 1499
DB 1441 TAAAGGTACAAATTTGACCAAAATCGATAAAACAAATAATACGTATCGTCAGACAATTTATGT 1500

QY 1500 CAATCAAGTGGAGATAACGTTATTCGCCGGTTTAAACAGGTAATTTAAACCAAAATAC 1559
DB 1501 CAATCAAGTGGAGATAACGTTATTCGCCGGTTTAAACAGGTAATTTAAACCAAAATAC 1560

QY 1560 GGATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTTAAAGTATATAAAGTAGATAA 1619
DB 1561 GGATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTTAAAGTATATAAAGTAGATAA 1620

QY 1620 TGCAGCTGATTTTCTGAAAGTACTTTGTGATCCAGAAACCTTTGAGGATGTCATAA 1679
DB 1621 TGCAGCTGATTTTCTGAAAGTACTTTGTGATCCAGAAACCTTTGAGGATGTCATAA 1680

QY 1680 TAGTGTGAATATTACATTTCCAAATCCAAATCAATTAAGTAGAGTTTAAATACGCGCTGA 1739
DB 1681 TAGTGTGAATATTACATTTCCAAATCCAAATCAATTAAGTAGAGTTTAAATACGCGCTGA 1740

QY 1740 TGATCAAAATTAACAACCGGTATATAGTAGTGTGTTAATGGTCAATTTGATCCGAATAGAA 1799

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

EP786519-A2.
30-JUL-1997.
07-JAN-1997; 97EP-00100117.
05-JAN-1996; 96US-0009861P.
(HUMA-) HUMAN GENOME SCI INC.
Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
MPI; 1997-374922/35.
Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
stored on computer readable medium and used in the production of anti-
S. aureus vaccines.
Claim 1; Page 1519-1521; 3271pp; English.
This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the *S. aureus* DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against *S. aureus* infection. The polypeptides can also be used in a kit for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the *S. aureus* DNA sequences contained on the computer readable medium
Sequence 4709 BP; 1657 A; 704 C; 908 G; 1313 T; 0 U; 127 Other;
Query Match 93.8%; Score 3279.6; DB 2; Length 4709;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 3346; Conservative 1; Mismatches 131; Indels 21; Gaps 4;
1 GGTACCATAAATTACATCTGCTTTTGAAGAAATATGATTTCAAGCTAGGATTACATTA 60
1204 GGTACCATAAATTACATCTGCTTTTGAAGAAATATGATTTCAAGCTAGGATTACATTA 1263
61 GGTAGAGTTCATATTAATAATAAATAATGTTTGCAATCAATCGTACGTTGCTGTTGTA 120
1264 GGTAGAGTTCATATTAATAATAAATAATGTTTGCAATCAATCGTACGTTGCTGTTGTA 1323
121 ATTCTTAAATAGCAATAATAAATAATGTTTGTAGTAAAGTATTAATTGTGGATATAAAA 180
1324 ATTCTTAAATAGCAATAATAAATAATGTTTGTAGTAAAGTATTAATTGTGGATATAAAA 1383
181 TATCGATACAAATTAATTGCTATTAATGCAATTTTGTAGTATAATTTCCATTAACAGAGATT 240
1384 TATCGATACAAATTAATTGCTATTAATGCAATTTTGTAGTATAATTTCCATTAACAGAGATT 1443
241 AATATATATC-TTAAAGGTTATATAGTTAATAATAAATGACCTTTTAAAGAGGGAATAA 299
1444 AATATATATCTTAAAGGTTATATAGTTAATAATAAATGACCTTTTAAAGAGGGAATAA 1503
300 AATCAATATAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 359
1504 AATGAATATAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1563

QY 360 AGTCTTGTAGTACGTTTAATCGGTTTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAG 419
DB 1564 AGTCTTGTAGTACGTTTAATCGGTTTTTGGACTACTCAGCAGTAAAGAGCAGATGCAAG 1623
QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAG 479
DB 1624 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTNN 1683
QY 480 TAGCGTTAGTCTGCACCTAAACAGAGGACACAAACGTGAGTGATATAAAACATGCTC 539
DB 1684 NNN 1743
QY 540 AACACTAATAATCGGAAACGAGTGTGGCGCAAAATCCAGCACAAACAGGAAACGACACA 599
DB 1744 AAACACTAATAATCGGAAACGAGTGTGGCGCAAAATCCAGCACAAACAGGAAACGACACA 1803
QY 600 ATCATCATCAACAAATCAACTACCGAAGAAACCGCGTAACTGCTGAAGCTACTACTAC 659
DB 1804 ATCATCATCAACAAATCAACTACCGAAGAAACCGCGTAACTGCTGAAGCTACTACTAC 1863
QY 560 GACAAAGCAATCAAGCTTAATACACCGCAACAACTCAATCAAGCAATCAAAATGCGGAGGA 719
DB 1864 GACAAAGCAATCAAGCTTAATACACCGCAACAACTCAATCAAGCAATCAAAATGCGGAGGA 1923
QY 720 ATTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATACTAATACAGTATCATCTGT 779
DB 1924 ATTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATACTAATACAGTATCATCTGT 1983
QY 780 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATATCTCAAC 839
DB 1984 AAATTCACCTCAAAATTTCTACAAATGCGGAAATGTTTCAACAAACGCAAGATATCTCAAC 2043
QY 840 TGAAGCAACACTTCAACCAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGA 899
DB 2044 TGAAGCAACACTTCAACCAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGA 2103
QY 900 TGTAAGTAAATCAACGCGTAAATCAAGTGGCGCTAGAAATGAGAGCATTTAGTTAGCGGC 959
DB 2104 TGTAAGTAAATCAACGCGTAAATCAAGTGGCGCTAGAAATGAGAGCATTTAGTTAGCGGC 2163
QY 960 AGTAGTGCAGATGCAACCGGAGCTGGCAGACAGATATTACGAATCAGTTGACGAATGTAC 1019
DB 2164 AGTAGTGCAGATGCAACCGGAGCTGGCAGACAGATATTACGAATCAGTTGACGAATGTAC 2223
QY 1020 AGTTGGTATTGACTCTGTACGACTGTATCCGACCAACAGAGGTTATGTCAAACTGAA 1079
DB 2224 AGTTGGTATTGACTCTGTACGACTGTATCCGACCAACAGAGGTTATGTCAAACTGAA 2283
QY 1080 TTATGGTTTTTTCAGTGCTTAATTCGTGTTTAAAGGTGACACATTCAAATAAATGCTGACC 1139
DB 2284 TTATGGTTTTTTCAGTGCTTAATTCGTGTTTAAAGGTGACACATTCAAATAAATGCTGACC 2343
QY 1140 TAAAGAAATTAACCTTAATGTTGTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
DB 2344 TAAAGAAATTAACCTTAATGTTGTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2403
QY 1200 AGATCAAGTATTCGCAATGCTGTAATCGATAGTGTAACTGTTATTTATATACATTTAC 1259
DB 2404 AGATCAAGTATTCGCAATGCTGTAATCGATAGTGTAACTGTTATTTATATACATTTAC 2463
QY 1260 AGACTATGTAATAACTAAAGATGATGTAAGCAACTTTTGACCATGCGCGCTTATATGA 1319
DB 2464 AGACTATGTAATAACTAAAGATGATGTAAGCAACTTTTGACCATGCGCGCTTATATGA 2523
QY 1320 CCTGAAATATGTTAAAGAGCAGTAATGTCATTTGGCTACTGGCATAGTAGTACAAAC 1379
DB 2524 CCTGAAATATGTTAAAGAGCAGTAATGTCATTTGGCTACTGGCATAGTAGTACAAAC 2583
QY 1380 AGCAAAACAAACAGTATTAGTAGATTATGAAATAATGTTAAAGTTTATTAACCTTATCTAT 1439
DB 2584 AGCAAAACAAACAGTATTAGTAGATTATGAAATAATGTTAAAGTTTATTAACCTTATCTAT 2643

1440 TAAAGGTACAAATTGACCAAAATCGATAAACAATAAATACGATATCGTCAGACAAATTTATGT 1499
1441 |||||
2644 TAAAGGTACAAATTGACCAAAATCGATAAACAATAAATACGATATCGTCAGACAAATTTATGT 2703
2645 |||||
1500 CAAATCAAGTGGAGATAAAGTTATTCGGCGGGTTTAAACAGGTAATTTAAACCAATATC 1559
1501 |||||
2704 CAAATCAAGTGGAGATAAAGTTATTCGGCGGGTTTAAACAGGTAATTTAAACCAATATC 2763
2705 |||||
1560 GGATAGTAATGCAATTAATAGATCAGACAAAATCAAGTATTTAAAGTATATAAAGTAGATATA 1619
1561 |||||
2764 GGATAGTAATGCAATTAATAGATCAGACAAAATCAAGTATTTAAAGTATATAAAGTAGATATA 2823
2765 |||||
1620 TGCAGCTGATTTATCTGAAAGTTACTTTGTGATCCAGAAAACCTTTGAGGATGTCCTACTAA 1679
1621 |||||
2824 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCCTACTAA 2883
2825 |||||
1680 TAGTGTGAATATTAATATTCGCAAAATCAAAATCAATTAAGTAGAGTTTAAATAGCGCTGA 1739
1681 |||||
2884 TAGTGTGAATATTAATATTCGCAAAATCAAAATCAATTAAGTAGAGTTTAAATAGCGCTGA 2893
2885 |||||
1740 TGATCAAAATTAACAACACCGTATATAGTAGTTGTTAATGGTCAATTTGATCCGAATAGCAA 1799
1741 |||||
2944 TGATCAAAATTAACAACCGTATATAGTAGTTGTTAATGGTCAATTTGATCCGAATAGCAA 3003
2945 |||||
1800 AGGTGATTTAGCTTTTACGTTCAACTTTTATATGGGTATAACTCGAATATATTTTGGCGCTC 1859
1801 |||||
3004 AGGTGATTTAGCTTTTACGTTCAACTTTTATATGGGTATAACTCGAATATATTTTGGCGCTC 3063
3005 |||||
1860 TATGTCATCGGACAAAGTAGTACATTTAATACGGATCAGGTTCTGGTACGCGTATCGA 1919
1861 |||||
3064 TATGTCATCGGACAAAGTAGTACATTTAATACGGATCAGGTTCTGGTACGCGTATCGA 3123
3065 |||||
1920 TAAACCAAGTTGTTCTGTAACAACCTGATGAGCTGTGTAATTTGAACCAATTTCCAGGGA 1979
1921 |||||
3124 TAAACCAAGTTGTTCTGTAACAACCTGATGAGCTGTGTAATTTGAACCAATTTCCAGGGA 3183
3125 |||||
1980 TTCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGATTTCTAATTCAGATAGCGGTTTCA 2039
1981 |||||
3184 TTCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGATTTCTAATTCAGATAGCGGTTTCA 3243
3185 |||||
2040 TTCGGGTAGTGATTTACATCAGATAGTGTTTCAATTCAGGAGTTCAGGAGTTCAGATTCAG 2099
2041 |||||
3244 TTCGGGTAGTGATTTACATCAGATAGTGTTTCAATTCAGGAGTTCAGGAGTTCAGATTCAG 3303
3245 |||||
2100 AAGTGATTCAGACTCAGCAGTGATTCAGATTCAGCAAGCGATTTCCGACTCAGCGAGCGA 2159
2101 |||||
3304 AAGTGATTCAGACTCAGCAGTGATTCAGATTCAGCAAGCGATTTCCGACTCAGCGAGCGA 3363
3305 |||||
2160 TTCGACTCAGACAAATGACTTCGATTCAGATAGCGATTTCTGATCAGACAGTGACTCAGA 2219
2161 |||||
3364 TTCGACTCAGACAAATGACTTCGATTCAGATAGCGATTTCTGATCAGACAGTGACTCAGA 3423
3365 |||||
2220 TTCGACAGTGACTCAGATTCAGATAGCGATTTCTGATCAGACAGTGACTCAGATTCAGA 2279
2221 |||||
3424 TTCGACAGTGACTCAGATTCAGATAGCGATTTCTGATCAGACAGTGACTCAGATTCAGAT 3483
3425 |||||
2280 TAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGATTTCCGACTCAGACAGCGA 2339
2281 |||||
3484 NNN 3543
3485 |||||
2340 TTCGACTCCGACAGTGATTTCCGACTCAGACAGCGATTCAGATTCGACAGTGATTTCCGA 2399
2341 |||||
3544 TTCGACTCCGACAGTGATTTCCGACTCAGACAGCGATTCAGATTCGACAGTGATTTCCGA 3603
3545 |||||
2400 CTCAGATAGCGATTTCCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA 2459
2401 |||||
3604 CTCAGATAGCGATTTCCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGA 3663
3605 |||||
2460 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGACTCAGATTCGACAGTGATA 2519
2461 |||||
3664 CAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGACTCAGATTCGACAGTGATA 3723
3665 |||||
2520 CTCGGATTCAGATAGCGATTCAGATTCGACAGTGACTCAGATTCGACAGTGACTCAGA 2579
2521 |||||

3724 CTCGGATTCAGATAGCGATTCAGATTCGACAGTGACTCAGATTCGACAGTGACTCAGA 3783
3725 |||||
2580 CTCAGACAGTGATTCGGATTCAGCGAGTGATTCGGATTCAGATAGTGATTCGGATTCGGA 2639
2581 |||||
3784 CTCAGACAGTGATTCGGATTCAGCGAGTGATTCGGATTCAGATAGTGATTCGGATTCGGA 3843
3785 |||||
2640 CAGTGACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 2699
2641 |||||
3844 CAGTGACTCGGATTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGA 3903
3845 |||||
2700 TTCGGACTCAGATAGCGATTCAGAAATCAGACAGCGATTCAGAAATCAGACAGCGATTCAGA 2759
2701 |||||
3904 TTCGGACTCAGAT-----AGCGATTCAGAAATCAGACAGCGATTCAGA 3945
3905 |||||
2760 TTCAGACAGCGACTCAGACAGTGACTCAGATTCAGATAGTGACTCGGATTCAGCGAGTGA 2819
2761 |||||
3946 TTCAGACAGCGACTCAGACAGTGACTCAGATTCAGATAGTGACTCGGATTCAGCGAGTGA 4005
3947 |||||
2820 TTCAGACTCAGGTAGTGACTCCGATTCATCAAGTGATTCGACTCAGAAAGTGATTCAAA 2879
2821 |||||
4006 TTCAGACTCAGGTAGTGACTCCGATTCATCAAGTGATTCGACTCAGAAAGTGATTCAAA 4065
4007 |||||
2880 TAGCGATTCGGATTCGATTAACAATATGTAGTTCGCGCTTAATTCACCTAAAAAATGG 2939
2881 |||||
4066 TAGCGATTCGGATTCGATTAACAATATGTAGTTCGCGCTTAATTCACCTAAAAAATGG 4125
4067 |||||
2940 TACTAATGCTTCTAATAAAAAATGAGCTAAGATAGTAAAGAACCAATTCAGGATACAGG 2999
2941 |||||
4126 TACTAATGCTTCTAATAAAAAATGAGCTAAGATAGTAAAGAACCAATTCAGGATACAGG 4185
4127 |||||
3000 TTCGAAAGTGAAGCAAAATACGTCACCTAATTTGGGATTTATTAGCATCAATAGTTTCATT 3059
3001 |||||
4186 TTCGAAAGTGAAGCAAAATACGTCACCTAATTTGGGATTTATTAGCATCAATAGTTTCATT 4245
4187 |||||
3060 ACTACTTTTCAGAGAAAAAAGAAAAATAGATAAGATAAAGATAAATGATAATGATTAATA 3119
3061 |||||
4246 ACTACTTTTCAGAGAAAAAAGAAAAATAGATAAAGATAAAGATAAATGATAATGATTAATA 4305
4247 |||||
3120 TTAATCATATGATTCATGAGAGCCCTTAAAGGTCCTTTTACTTCGATTTTCC 3179
3121 |||||
4306 TTAATCATATGATTCATGAGAGNA-ACRCCTTAAAGGTCGCTTTTACTTCGATTTTCC 4364
4307 |||||
3180 AAATATATTTGTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3239
3181 |||||
4365 AAATATATTTGTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4424
4366 |||||
3240 GTAGATGTTTATATAATTTGGCTTGGCGAAAAAATAGGGTGTAAAGGTAGGTTGTTAATTAG 3299
3241 |||||
4425 GTAGATGTTTATATAATTTGGCTTGGCGAAAAAATAGGGTGTAAAGGTAGGTTGTTAATTAG 4484
4426 |||||
3300 GGAAAAATTAAGGAGAAAAATACAGTTGAAAAATTAATTTGCTAGTTTATCATTTGGGAGCAT 3359
3301 |||||
4485 GGAAAAATTAAGGAGAAAAATACAGTTGAAAAATTAATTTGCTAGTTTATCATTTGGGAGCAT 4544
4486 |||||
3360 TATGTGATACAAAATTTGGGAAAGTAAATCGTCGAGTGACGTTTCTGGGGAAGA 3419
3361 |||||
4545 TATGTGATACAAAATTTGGGAAAGTAAATCGTCGAGTGACGTTTCTGGGGAAGA 4604
4546 |||||
3420 ATCCATATGATCTGAGTCGTTGAAATCGACTAATAATAAAAAAATAAATCTAGAACAGTAG 3479
3421 |||||
4605 ATCCATATGATCT-AGTCGTTGAAATCGACTAATAATAAAAAAATAAATCTAGAACAGTAG 4663
4606 |||||
3480 AAGAGTATTAAGAAAAGCTT 3498
3481 |||||
4664 AAGAGTATTAAGAAAAGAT 4682
4665 |||||

RESULT 4
ABT14896
ID ABT14896 standard; DNA; 2799 BP.
XX
AC ABT14896;

```
XX 06-MAR-2003 (first entry)
DT Pathogen specific antigen related staphylococcal DNA SEQ ID No 41.
DE
DE
DE
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
XX hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW autoimmune disease; HIV; hepatitis; gene; ds.
XX
OS Staphylococcus sp.
XX
XX WO200259148-A2.
XX
XX 01-AUG-2002.
XX
XX 21-JAN-2002; 2002WO-EP000546.
XX
XX 26-JAN-2001; 2001AT-00000130.
XX
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
XX Meinke A, Nagy E, Von Ahlsen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vykvytska O, Etz H, Dryia A, Weichhart T, Hafner M;
PI Tempelmaier B;
XX
XX WPI; 2003-075410/07.
XX
XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.
XX
XX Example 7; Page 151; 252pp; English.
XX
XX The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX allergen, a tissue or host prone to auto-immunity, where the antigens are
XX used in a vaccine, comprises providing antibody preparation from a plasma
XX pool of a type of animal, or individual sera with antibodies against the
XX specific pathogen, tumour, allergen, tissue or host prone to auto-
XX immunity. The hyperimmune serum-reactive antigens comprising any of the
XX 62 sequences of 53-2261 amino acids fully defined in the specification,
XX or their hyperimmune fragments are useful for the manufacture of a
XX pharmaceutical preparation, particularly a vaccine against staphylococcal
XX infections or colonisation against S. aureus or S. epidermidis. The
XX preparation of antibodies is useful for the manufacture of a medicament
XX for treating or preventing staphylococcal infections or colonisation
XX against S. aureus or S. epidermidis. The antibody preparations may also
XX be used for diagnostic and imaging purposes. Other conditions that can be
XX treated include cancer, autoimmune diseases or infections caused by viral
XX (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX polynucleotide sequence represents staphylococcal DNA relating to the
XX method for identifying and producing pathogen specific antigens of the
XX invention
XX
XX Sequence 2799 BP; 981 A; 527 C; 582 G; 709 T; 0 U; 0 Other;
XX
XX Query Match 80.0%; Score 2797.4; DB 8; Length 2799;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2798; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 301 ATGAATATGAAGAAAAAGAAAAACACGCAATTCGGAATAAATCGATTGGGTGGCTTCA 360
XX 1 ATGAATATGAAGAAAAAGAAAAACACGCAATTCGGAATAAATCGATTGGGTGGCTTCA 60
XX
XX 361 GTGCTTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT 420
XX 61 GTGCTTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT 120
XX
XX 421 GAAATATGCTTACGCAATCTGATAGCGCAAGTAACGAACGCAAAAGTAATGATTCAGT 480
XX 121 GAAATATGCTTACGCAATCTGATAGCGCAAGTAACGAACGCAAAAGTAATGATTCAGT 180
```

1561 GATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAAT 1620
1261 GATAGTAATGCAATTAATAGATCAGCAAAATACAAAGTATTAAAGTATATAAAGTAGATAAT 1320
1621 GAGCTGATTTATCTGAAAGTTACTTTGTGTAATCCAGAAAATTGAGGATGTCATAAT 1680
1321 GAGCTGATTTATCTGAAAGTTACTTTGTGTAATCCAGAAAATTGAGGATGTCATAAT 1380
1681 AGTGTCAATATTACATTTCCCAATCCAAATCAATATAAAGTATAGATTTAATAGCGCTGAT 1740
1381 AGTGTCAATATTACATTTCCCAATCCAAATCAATATAAAGTATAGATTTAATAGCGCTGAT 1440
1741 GATCAAAATTACAAACCGGATATAGATTTGTTAAATGGTCAATATTAATGATCCGAATACAAA 1800
1441 GATCAAAATTACAAACCGGATATAGATTTGTTAAATGGTCAATATTAATGATCCGAATACAAA 1500
1801 GGTGATTTAGCTTTCAGTTCACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT 1860
1501 GGTGATTTAGCTTTCAGTTCACTTTATATGGGTATAAATCGAATATAATTTGGCGCTCT 1560
1861 ATGTCATGGGACAAACGAAGTATGATTTAATAGCGATCAGTTCTGGTGACGGTATCGAT 1920
1561 ATGTCATGGGACAAACGAAGTATGATTTAATAGCGATCAGTTCTGGTGACGGTATCGAT 1620
1921 AAACCAAGTTGTTCTCGAACCAACCTGATGAGCCTGGTGAATTTGAACCAATTCAGAGGAT 1980
1621 AAACCAAGTTGTTCTCGAACCAACCTGATGAGCCTGGTGAATTTGAACCAATTCAGAGGAT 1680
1981 TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAAATTCAGATAGCGGTTTCAGAT 1740
1681 TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAAATTCAGATAGCGGTTTCAGAT 1740
2041 TCGGGTAGTGATTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCA 2100
1741 TCGGGTAGTGATTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCA 1800
2101 AGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 2160
1801 AGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 1860
2161 TCGGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 2220
1861 TCGGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 1920
2221 TCGGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 2280
1921 TCGGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGAT 1980
2281 AGCGATTCAGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAG 2340
1981 AGCGATTCAGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAGTGATTCAGATTCAGCGAG 2040
2341 TGTGATTCGACAGTATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2400
2041 TGTGATTCGACAGTATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2100
2401 TCAGATAGCGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2460
2101 TCAGATAGCGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2160
2461 AGCGATTCAGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2520
2161 AGCGATTCAGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2220
2521 TCGGATTCAGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2580
2221 TCGGATTCAGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2280
2581 TCAGACAGTATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2640
2281 TCAGACAGTATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2340
2641 AGTGACTCGGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2700

2341 AGTGACTCGGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2400
2701 TCGGATTCAGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2760
2401 TCGGATTCAGATTCAGATTCAGCGAGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2460
2761 TCAGACAGCGATTCAGACAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2820
2461 TCAGACAGCGATTCAGACAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2520
2821 TCAGACTCAGGTAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2880
2521 TCAGACTCAGGTAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2580
2881 AGCGATTCGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2940
2581 AGCGATTCGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2640
2941 ACTAATGCTTCTTAATAAATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3000
2641 ACTAATGCTTCTTAATAAATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2700
3001 TCTGAAGATGAAGCAAAATACGTCACATAATTTGGGGATTTAGCATCAATAGATTCATTA 3060
2701 TCTGAAGATGAAGCAAAATACGTCACATAATTTGGGGATTTAGCATCAATAGATTCATTA 2760
3061 CTACTTTTTCAGAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 3099
2761 CTACTTTTTCAGAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 2799

RESULT 5

ACF73781
ID ACF73781 standard; DNA; 2781 BP.

XX ACF73781;

XX 20-NOV-2003 (first entry)

XX Staphylococcus aureus DNA #1461.

XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target; gene; ds.

XX Staphylococcus aureus.

XX WO200294868-A2.

XX 28-NOV-2002.

XX 27-MAR-2002; 2002WO-IB002637.

XX 27-MAR-2001; 2001GB-00007661.

XX (CHIR-) CHIRON SPA.

XX Masignani V, Mora M, Scarselli M;

XX WPI; 2003-120786/11.

XX P-PSDB; ABM72221.

XX New Staphylococcus aureus protein, useful as a vaccine for treating or
preventing Staphylococcal infection, specifically an infection caused by
S. aureus, e.g. sepsis.

XX Claim 6; SEQ ID NO 2921; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
Staphylococcus aureus. A composition comprising the S. aureus protein, a
nucleic acid encoding the protein, or an antibody to the protein, is
useful as a pharmaceutical, particularly as a vaccine for treating or
preventing infection due to Staphylococcus bacteria, specifically an

CC infection caused by *S. aureus*. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel *S. aureus* genes of the invention
 XX
 SQ

Query Match	78.5%;	Score 2746.6;	DB 8;	Length 2781;	
Best Local Similarity	99.2%;	Pred. No. 0;			
Matches 2777;	Conservative 0;	Mismatches 4;	Indels 18;	Gaps 1;	
301	ATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTGCCTCA	360			
Db	1 ATGAATATGAAGAAAAAGAAAAACAGCAATTCGAAAAAATCGATTGGCGTGCCTCA	60			
361	GTGCTTAGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT	420			
Db	61 GTGCTTAGTAGTACGTTAAATCGGTTTGGACTACTCAGCAGTAAGAAGCAGATGCAAGT	120			
421	GAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAGCAAAAGTAATGATTCAAGT	480			
Db	121 GAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAGCAAAAGTAATGATTCAAGT	180			
481	AGCGTTAGTGTGCACCTTAAACACAGACGACACAAACGTTAGTGTACTTAAACATCGTCA	540			
Db	181 AGCGTTAGTGTGCACCTTAAACACAGACGACACAAACGTTAGTGTACTTAAACATCGTCA	240			
541	AACACTAATATGCGGAAACGAGTGTGGCGAAATCCAGCACACAGGAAGACACACAA	600			
Db	241 AACACTAATATGCGGAAACGAGTGTGGCGAAATCCAGCACACAGGAAGACACACAA	300			
601	TCATCATCAAAATGCAACTTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTACG	660			
Db	301 TCATCATCAAAATGCAACTTACGGAAGAAACGCGGTAACTGGTGAAGTACTACTACG	360			
661	ACAAACGAATCAAGTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA	720			
Db	361 ACAAACGAATCAAGTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA	420			
721	TTAGTGAATCAAAACAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA	780			
Db	421 TTAGTGAATCAAAACAGTAATGAACGACTTTTAAATGATCTAATACAGTATCATCTGTA	480			
781	AATTCACTCAAAATCTCAAAATGCGGAAATGTTTCAACAAACGAGATACATCAACT	840			
Db	481 AATTCACTCAAAATCTCAAAATGCGGAAATGTTTCAACAAACGAGATACATCAACT	540			
841	GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT	900			
Db	541 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT	600			
901	GTAGTTAATCAAGCGGTTAATACAGTCCGCTAGATGAGAGCATTTAGTTTACGGCA	960			
Db	601 GTAGTTAATCAAGCGGTTAATACAGTCCGCTAGATGAGAGCATTTAGTTTACGGCA	660			
961	GTAGCTGAGATGACCGGAGCTGGGACAGATATTACGAATCAGTTGACGAATGTGACA	1020			
Db	661 GTAGCTGAGATGACCGGAGCTGGGACAGATATTACGAATCAGTTGACGAATGTGACA	720			
1021	GTGGTATTGACTCTGGTACAGTGTGTATCCGCAACGAGGTTATGTCAAACTGAAAT	1080			
Db	721 GTGGTATTGACTCTGGTACAGTGTGTATCCGCAACGAGGTTATGTCAAACTGAAAT	780			
1081	TATGGTTTTTCAGTGCCTTAATCTGCTGTTAAAGTGACACATTCAAAATACCTGTACCT	1140			
Db	781 TATGGTTTTTCAGTGCCTTAATCTGCTGTTAAAGTGACACATTCAAAATACCTGTACCT	840			
1141	AAAGAAATTAACCTTAAATGGTGAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA	1200			
Db	841 AAAGAAATTAACCTTAAATGGTGAACCTTCAACTGCTAAAGTGCCACCAATTTATGGCTGGA	900			
1201	GATCAAGTATGGCAAAATGGTGAATCGATAGTGAATGTTATTATACATTACA	1260			

Db	901	GATCAAGTATTTGGCAAAATGGTGAATCGATAGTGAATGTTATTATACATTACA	960
Qy	1261	GACTATCTAATACTAAGATGATGTAAGAACCACTTTGACCATTGCGCTTATATTGAC	1320
Db	961	GACTATGTAATACTAAGATGATGTAAGAACCACTTTGACCATTGCGCTTATATTGAC	1020
Qy	1321	CCTGAAAATGTTAAAAAGACAGGTAATGTGCATTTGGCTACTGGCATAGGTAGTACAACA	1380
Db	1021	CCTGAAAATGTTAAAAAGACAGGTAATGTGCATTTGGCTACTGGCATAGGTAGTACAACA	1080
Qy	1381	GCAAAACAAACAGATTATTAGTAGATTATGAAAAATATGGTAAGTTTATACTTCTATT	1440
Db	1081	GCAAAACAAACAGATTATTAGTAGATTATGAAAAATATGGTAAGTTTATACTTCTATT	1140
Qy	1441	AAAGGTCAATTTGACCAAAATCGATAAAACAAATTAATACGTATTCGTACAGCAATTTATGTC	1500
Db	1141	AAAGGTCAATTTGACCAAAATCGATAAAACAAATTAATACGTATTCGTACAGCAATTTATGTC	1200
Qy	1501	AATCCAAAGTGAGATAACGTTATTGCGCCGGTTTAAACAGGTAATTTAAAAACCAATACG	1560
Db	1201	AATCCAAAGTGAGATAACGTTATTGCGCCGGTTTAAACAGGTAATTTAAAAACCAATACG	1260
Qy	1561	GATAGTAATGCAATTAATAGATCAGCAAAATAACAAGTATTTAAAGTATATAAAGTAGATAAT	1620
Db	1261	GATAGTAATGCAATTAATAGATCAGCAAAATAACAAGTATTTAAAGTATATAAAGTAGATAAT	1320
Qy	1621	GCAGCTGATTTATCTGAAAGTACTTTGTGAATCCAGAAACTTTGAGGAGTGTCACTAAT	1680
Db	1321	GCAGCTGATTTATCTGAAAGTACTTTGTGAATCCAGAAACTTTGAGGAGTGTCACTAAT	1380
Qy	1681	AGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACGCCCTGAT	1740
Db	1381	AGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGAGTTTAAATACGCCCTGAT	1440
Qy	1741	GATCAAAATTAACAACCGGTATATAGTAGTTTAAATGCTCATATTTGATCCGAATAGCAAA	1800
Db	1441	GATCAAAATTAACAACCGGTATATAGTAGTTTAAATGCTCATATTTGATCCGAATAGCAAA	1500
Qy	1801	GGTGATTTAGCTTTACGTTCAACTTTATATGGGTATTAATCGAATATAATTTGGCGCTCT	1860
Db	1501	GGTGATTTAGCTTTACGTTCAACTTTATATGGGTATTAATCGAATATAATTTGGCGCTCT	1560
Qy	1861	ATGTCATGGGACAAACGAAGTAGCAATTTAATAACGGATCAGGTTCTGGTGACGGTATCGAT	1920
Db	1561	ATGTCATGGGACAAACGAAGTAGCAATTTAATAACGGATCAGGTTCTGGTGACGGTATCGAT	1620
Qy	1921	AAACAGGTTGTTCTGTGAACACCTGATGAGCTGGTGAAATTTGAACCAATTTCCAGAGGAT	1980
Db	1621	AAACAGGTTGTTCTGTGAACACCTGATGAGCTGGTGAAATTTGAACCAATTTCCAGAGGAT	1680
Qy	1981	TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTTAATTCAGATAGCGGTTTCAGAT	2040
Db	1681	TCAGATTTGACCCAGGTTTCAGATTTCTGGCAGCGATTCTTAATTCAGATAGCGGTTTCAGAT	1740
Qy	2041	TCGGGTAGTGAATTTACATCAGATPAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA	2100
Db	1741	TCGGGTAGTGAATTTCTCATCAGATPAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA	1800
Qy	2101	AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCAACGATTCGGACTCAGCGAGGAT	2160
Db	1801	AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCAACGATTCGGACTCAGCGAGGAT	1860
Qy	2161	TCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGAT	2220
Db	1861	TCCGACTCAGACAAATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGAT	1920
Qy	2221	TCCGACAGTGAATTCAGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGATTCAGAT	2280
Db	1921	TCCGACAGTGAATTCAGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGATTCAGAT	1980
Qy	2281	AGCGATTCAGATTCAGATAGCGATTCTGACTCAGACAGTGAATTCGGACTCAGACAGGAT	2340

QY 481 AGCGTTAGTGTGCTACCTAAAAACAGACGACACAAAACGTGAGTGATCTATAAAACATCGTCA 540
DB 181 AGCGTTAGTGTGCTACCTAAAAACAGACGACACAAAACGTGAGTGATCTATAAAACATCGTCA 240
QY 541 AACACTAATATGCGGAAACAGAGTGTGGCGGAAATCCAGGACACAGGAAACGACAAA 600
DB 241 AACACTAATATGCGGAAACAGAGTGTGGCGGAAATCCAGGACACAGGAAACGACAAA 300
QY 601 TCATCATCAACAAATGCACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTACG 660
DB 301 TCATCATCAACAAATGCACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTACG 360
QY 661 ACAACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGAA 720
DB 361 ACAACGAATCAAGCTAATAACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGAA 420
QY 721 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATACCTAAACAGTATCACTCTGTA 780
DB 421 TTAGTGAATCAACAAAGTAATGAACGACTTTTAAATGATACCTAAACAGTATCACTCTGTA 480
QY 781 AATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACACGCAAGATACCTCAACT 840
DB 481 AATTCACCTCAAAATTTCTACAAATCGGAAATGTTTCAACACGCAAGATACCTCAACT 540
QY 841 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATCCAAAGTAATAAGAT 900
DB 541 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATCCAAAGTAATAAGAT 600
QY 901 GTAGTTAATCAAGCGGTAAATACAGTGGCGCTAGAGTAAAGTACAGTAAAGTACAGTAA 960
DB 601 GTAGTTAATCAAGCGGTAAATACAGTGGCGCTAGAGTAAAGTACAGTAAAGTACAGTAA 660
QY 961 GTAGCTGAGATGACCGGACGCTGGCACAGATATTTACGAATCAGTTGACGAATGTGACA 1020
DB 661 GTAGCTGAGATGACCGGACGCTGGCACAGATATTTACGAATCAGTTGACGAATGTGAAA 720
QY 1021 GTTGTATGACTCTGGTACGACTGTGTATCGGACCAAGCAGGTTATGTCAAACTGAAT 1080
DB 721 GTTGTATGACTCTGGTACGACTGTGTATCGGACCAAGCAGGTTATGTCAAACTGAAT 780
QY 1081 TATGTTTTTTCAGTGCCTAAATCTCTGTGTTAAAGTGACACANTTCAAAATACCTGACT 1140
DB 781 TATGTTTTTTCAGTGCCTAAATCTCTGTGTTAAAGTGACACANTTCAAAATACCTGACT 840
QY 1141 AAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTCCACCAATATGGCTGGA 1200
DB 841 AAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTCCACCAATATGGCTGGA 900
QY 1201 GATCAAGTATGGCAAAATGGTGTAAATCGATAGTGTAAATGTTTATTTATCAATTTACA 1260
DB 901 GATCAAGTATGGCAAAATGGTGTAAATCGATAGTGTAAATGTTTATTTATCAATTTACA 960
QY 1261 GACTATGTAAATTAAGTAAAGTAAAGCAACTTTGACCATGCCGCTTATATTGAC 1320
DB 961 GACTATGTAAATTAAGTAAAGTAAAGCAACTTTGACCATGCCGCTTATATTGAC 1020
QY 1321 CCGTGAATATGTTAAAAACAGAGTAAATGTGACATTTGGCTACTGGCATAGGTAGTACAA 1380
DB 1021 CCGTGAATATGTTAAAAACAGAGTAAATGTGACATTTGGCACTGGCATAGGAACCAATACT 1080
QY 1381 GCAAAACAAACAGTATTTAGTAGATTTAGAAAATATGGTAAATTTTAACTTATCTATT 1440
DB 1081 GCTAGTAAGACAGTATTTAATCGACTATGAGAAATATGGAACAATTCCAATTTATCAATTT 1140
QY 1441 AAAGGTACAAATTTGACCAATCGATAAAACAAATATACGTATCGTCAGACAAATTTATGTC 1500
DB 1141 AAAGGTACAAATTTGACCAATCGATAAAACAAATATACGTATCGTCAGACAAATTTATGTC 1200
QY 1501 AATCAAGTGGAGATPAACGTTATTCGCGGTTTTTAAACAGGTAATTTAAAAACCAATAG 1560
DB 1201 AATCAAGCGGAGATAACGTTGTGTTTACCTGCTTAAACAGGTAATTTTAAATCCTAATACA 1260
QY 1561 GATAGTAATGCATTAATAGATCAGCAAAATACAAAGTATTAAGTATATAAGTAGATAAT 1620

DB 1261 AAGACTAATCGCTTAATAGATGCAAAAAACACTGATATTAAGTTTATAGAGTCGATAAT 1320
QY 1621 GCAGCTGATTTATCTGAAAGTACTTTCTGTAATCCAGAAAACCTTTGAGGATCTCACTAAT 1680
DB 1321 GCTAATGATTTATCTGAAAGTATTTATGTAATCTAGCGATTTTGAAGATTAATAAT 1380
QY 1681 AGTGTGAATATTTACATTCCTCAAAATCCAAATCAATATAAAGTAGAGTTTAAATCGGCTGAT 1740
DB 1381 CAAAGTGAATTTTCAATTTCCAAATGCTAATCAATACAAAGTAGAATTTCTTACGGACGAT 1440
QY 1741 GATCAAAATTACAAACCGGTATATAGTGTGTTAATGGTCAATATGATTCGGAATAGCAAAA 1800
DB 1441 GATCAAAATTACAAACCGGTATATAGTGTGTTAATGGTCAATATGATTCGGAATAGCAAAA 1500
QY 1801 GGTGATTTAGCTTTTACGTTCAACTTTATATGGGTATAAATCGAATATAATTTTGGCGCTCT 1860
DB 1501 GGTGATTTAGCACTACGTTTCGACATTTTATGGTATGATTTCTAAATTTATATGGAGATCT 1560
QY 1861 ATGTCTATGGGACAAACGAAGTAGCATTTTAAATACGGATCAGGTTCTGGTGACGATTCGAT 1920
DB 1561 ATGTCTATGGGACAAACGAAGTAGCATTTTAAATACGGATCAGGTTCTGGTGACGATTCGAT 1620
QY 1921 AAACCAAGTTGTTCTGTAACCACTGATGAGCCTGGTGAAATTTGAACCAATTTCCAGAGAT 1980
DB 1621 AAACCAAGTTGTTCTGTAACCACTGATGAGCCTGGTGAAATTTGAACCAATTTCCAGAGAT 1680
QY 1981 TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGAGCGATTTCTAATTCAGATAGCGGTTTCAGAT 2040
DB 1681 TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGAGCGATTTCTAATTCAGATAGCGGTTTCAGAT 1740
QY 2041 TCGGTAGTGTATTTACATCAGATAGTGTGTTTCAGATTCAGCGAGTGTATTCAGATTCAGCA 2100
DB 1741 TCTGSCAGTGTATTTACATCAGATAGTGTGTTTCAGATTCAGCGAGTGTATTCAGATTCAGCA 1800
QY 2101 AGTGTATTCAGACTCAGCGAGTGTATTCAGATTCAGCAAGCGATTCGAGTCCAGCGAGGAT 2160
DB 1801 AGTGTATTCAGACTCAGCGAGTGTATTCAGATTCAGCAAGCGATTCAGATTCAGCAAGTGTAT 1860
QY 2161 TCCGACTCAGACAAATGACTTCGATTCAGATAGCGATTTCTGACTCAGACAGTCACTCAGAT 2220
DB 1861 TCAGATTCAGCAAGTGTATTCAGACTCAGCAAGTGTATTCAGATTCAGCAAGTGTATTCAGAT 1920
QY 2221 TCCGACAGTGTATTCAGATTCAGATAGCGATTTCTGACTCAGACAGTGTACTCAGATTCAGAT 2280
DB 1921 TCAGCAAGCGATTCAGATTCAGCGAGCGATTCAGATTCAGCGAGCGATTCAGATTCAGCG 1980
QY 2281 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGTATTCGACTCAGACAGCGAT 2340
DB 1981 AGTGTATTCGACTCAGCGAGCGATTCAGACTCAGATAGTGTACTCAGATTCGATAGCGAT 2040
QY 2341 TCTGACTCCGACAGTGTATTCGACTCAGACAGCGATTCAGATTCGCGACAGTGTACTCAGAT 2400
DB 2041 TCCGACTCAGATAGCGATTCAGATTCAGACAGCGATTTCTGACTCAGACAGCGATTTCTGAC 2100
QY 2401 TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGAC 2460
DB 2101 TCAGACAGTGTACTCAGATTCGCGATTCGATAGCGATTCGACTCAGACAGTGTACTCAGATTCGAT 2160
QY 2461 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCGACAGTGTACTCAGATTCGCGACAGTGTAC 2520
DB 2161 AGCGATTCAGATTCAGACAGTGTATTCAGACTCAGATAGCGATTCAGATTCGCGACAGTGTAC 2220
QY 2521 TCGGATTCAGATAGCGATTCAGATTCGCGACAGTGTACTCAGATTCGCGACAGTGTACTCAGAC 2580
DB 2221 TCAGACTCAGACAGCGATTCAGATTCGCGATAGCGATTCAGATTCGCGACAGTGTACTCAGAT 2280
QY 2581 TCAGACAGTGTATTCGATTCAGCGAGTGTATTCGAGATTCAGATTCAGATTCGAGATTCGAGATTCGAG 2640
DB 2281 TCCGATAGTGTACTCAGATTCAGCGAGTGTATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2340
QY 2641 AGTGTATTCGATTCAGATAGCGACTCAGATTCGCGATAGCGATTCGAGTTCGATTCAGATTCAGATTCGAT 2700

QY	1021	GTTCGGTATTGACTCTCGTACGACTGTGTATCCGACACCAAGCAGGTTATGTCAAACTGAAT	1080
DB	721	GTCTACTATTGACTCTCGTACGACTGTGTATCCGACACCAAGCAGGTTATGTCAAACTGAAT	780
QY	1081	TATGGTTTTTTCAGTGCCTTAATTCCTGCTGTTTAAAGGTGACACATTCAAAAATAACTGTACCT	1140
DB	781	TATGGTTTTTTCAGTGCCTTAATTCCTGCTGTTTAAAGGTGACACATTCAAAAATAACTGTACCT	840
QY	1141	AAAGAAATAAACTTTAAATGGTGTAACTTCAACTGCTAAAGTGCCACCAATATATGGCTGGA	1200
DB	841	AAAGAAATAAACTTTAAATGGTGTAACTTCAACTGCTAAAGTGCCACCAATATATGGCTGGA	900
QY	1201	GATCAAGTATTTCGGCAAAATGGTGAATCGATAGTGCATGTAATGTTATTTATACATTTACA	1260
DB	901	GATCAAGTATTTCGGCAAAATGGTGAATCGATAGTGCATGTAATGTTATTTATACATTTACA	960
QY	1261	GACTATGTAATACTATAAGATGATGTAAAAAGCAACTTTTGACCATGCGCGCTTATATTGAC	1320
DB	961	GACTATGTTGATACTAAAGAAAAATGTAAACAGCTAAATATTACTATGCCAGCTTATATTGAC	1020
QY	1321	CCTGAAATGTTAAAAAGACAGGTAATGTGACATTTGCTACTGCTACTGCGATAGGTAGTACAA	1380
DB	1021	CCTGAAATGTTAAAAAGACAGGTAATGTGACATTTGCTACTGCTACTGCGATAGGTAGTACAA	1080
QY	1381	GC AAA CAAA CAGTATTAGTAGATTATGAAAAATATGTTAAGTTTATAACTTATCTATT	1440
DB	1081	GCTAGTAGACAGTATTAAATCGACTATGAGAAATATGGCAAAATTTCCATAATTTATCAATT	1140
QY	1441	AAAGGTACAAATGACAAATTCGATATAAACAAAATAATACGTATCGTCAGACAAATTTATGTC	1500
DB	1141	AAAGGTACAAATGATCAAAATCGATATAAACAAAATAATACGTATCGCCAAA CAAATTTATGTC	1200
QY	1501	AATCCAAAGTCGAGATAACGTTATTTCGCCCGGTTTTTAACAGGTAAATTTAAACCAAATAAG	1560
DB	1201	AATCCAAAGCGGAGATAACGTTATTTCGCCCGTAAACAGGTAAATTTTAAATTCCTATACA	1260
QY	1561	GATAGTAATGCAATTAATPAGATACAGCAAAAATACAAGTATTAAGGTATATAAAGTAGATAAT	1620
DB	1261	AAGAGTAATGCGTTAAATAGATGCAAAAAACACTGATTTAAAGTTTATAGTAGTAGATAAT	1320
QY	1621	GCAGCTGATTTTACTGAAAGTACTTTGTGTAATCCAGAAAACCTTTGAGGATGTCACTAAT	1680
DB	1321	GCTAATGATTTATCTGAAAGTATTATGTGTAATCCTAGCGGATTTTGAAGATGTAACCTAAT	1380
QY	1681	AGTGTGAATATTACATTTCCCAAAATCCAAATCAATATAAAGTAGAGTTTAAATAGCCCTGAT	1740
DB	1381	CAAGTTAGAAATTTCAATTTCCAAATGCTAATCAATACAAAGTAGAAATTTCTCTACGAGCAT	1440
QY	1741	GATCAAAATTACAACCGGTATATAGTAGTTGTTAAATGGTCAATTTGATCCGAATAGCAAA	1800
DB	1441	GATCAAAATTACAACCGGTATATGTTAGTTGTTTAAATGGCCATATTTGATCCTGCTAGCACA	1500
QY	1801	GGTGATTTAGCTTTTACGTTCCAACTTTATATGGGTATAAATCCGAATATAAATTTGGCGGCTCT	1860
DB	1501	GGTGATTTAGCACTACGTTCCGACATTTTATGGTTATGATTTCTAATTTTATATCGAGATCT	1560
QY	1861	ATGTCATGGGACAAACGAAGTAGCATTTAATTAACGGATCAGGTTCTGGTGACGTTATCGAT	1920
DB	1561	ATGTCATGGGACAAACGAAGTAGCATTTAATAACGGATCAGGTTCTGGTGACGTTATCGAT	1620
QY	1921	AAACCAAGTTGTTCTGTAAACCACTGTATGACGCTGGTGAAATTCGAACTTCCAGAGGAT	1980
DB	1621	AAACCTGTTGTTCTGTAAACCACTGTATGACCTGGTGAAATTTGAACCAATTCAGAGGAT	1680
QY	1981	TCAGATTTCTGACCCAGGTTTCAGATTTCTGGCAGCGAATTTCTAATTTTCAGATAGCGGTTTCAGAT	2040
DB	1681	TCAGATTTCTGACCCAGGTTTCAGATTAGTGGTTTCAGATTTCTAATTTTCAGATAGCGGTTTCAGAT	1740
QY	2041	TCGGGTAGTCAATCTCAATCAGATAGTGGTTTCAGATTCACGAGTGTATTCAGATTCAGCA	2100
DB	1741	TCGGGTAGTCAATCTCAATCAGATAGTGGTTTCAGATTCACGAGTGTATTCAGATTCAGCG	1800
QY	2101	AGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGAATTCGCACTCAGCGAGCGAT	2160

[illegible]

Db 2881 CGCCCTAATTCACCTAAATGCTACTAATGCTTCTAATAAAAAATGAGGCTAAAGATAGT 2940
QY 2977 AAGAACCATTTACAGATACAGGTTCTGAAGATGAAGCAATACGTCACCTAAATTTGGGGA 3036
Db 2941 AAGAGCCCATTTACAGATACAGGTTCTGAAGGTAAGCGAATACGTCACCTAAATTTGGGGA 3000
QY 3037 TTATTAGCATCAATAGGTTTCATTACTACTTTTTCAGAGGAAAAAAGAAAAATAAGATAAG 3096
Db 3001 TTATTAGCATCAATAGGTTTCATTACTACTTTTTCAGAGGAAAAAAGAAAAATAAGATAAG 3060
QY 3097 AATAA 3102
Db 3061 AATAA 3066

RESULT 8

AAS51834

ID AAS51834 standard; DNA; 3063 BP.

XX AAS51834;

XX 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #251.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-61145770.

XX P-PSDB; AAU33975.

XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 4416; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence encodes an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 3063 BP; 1056 A; 598 C; 638 G; 771 T; 0 U; 0 Other;
SQ Query Match 59.9%; Score 2093.8; DB 4; Length 3063;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 2542; Conservative 0; Mismatches 257; Indels 264; Gaps 3;
QY 301 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAATAATCGATTCGCGTGGCTTCA 360
Db 1 ATGAATATGAAGAAAAAAGAAAAACACGCAATTCGGAATAATCGATTCGCGTGGCTTCA 60
QY 361 GTGCTTGTAGGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAGTAAAGAGCAGATGCAAGT 420
Db 61 GTGCTTGTAGGTACGTTAATCGGTTTTTGGACTACTCAGCAGTAGTAAAGAGCAGATGCAAGT 120
QY 421 GAAATAGTGTTCAGCAATCTGTAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAAGT 480
Db 121 GAAACACAGTGTACGCAATCCGATAGCGCAAGTAACGAAAGCAAAAGTAGTATTCAAGT 180
QY 481 AGCGTTAGTGTCCACCTTAAACACAGACGACAAAAGTGTAGTACTATAAAACATGTGTC 540
Db 181 AGCGTTAATGCTGCACCTTAAACACAGACGACAAAAGTGTAGTACTATAAAACATGTGTC 240
QY 541 AACACTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACACACAGGAAACGACAAA 600
Db 241 AACACTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCACACACAGGAAACGACAAA 300
QY 601 TCATCATCAACAAATGCAACTACCGAAGAAACCCCGGTAACTGGTGAAGCTACTTACTACG 660
Db 301 TCAGCATCAACAAATGCAACTACCGAAGAAACCCCGGTAACTGGTGAAGCTACTTACTACG 360
QY 661 ACAACGAATCAAGCTAATACACCGGCAACACACTCAATCAAGCAATCAAAATCGGAGGAA 720
Db 361 GCAACGAAGCAAGCTAATACACCGGCAACACACTCAATCAAGCAATCAAAATCGGAGGAA 420
QY 721 TTAGTGAATCAAAACAAAGTAATGAACGACTTTTAAATGATACTAATACAGTATCATCTGTA 780
Db 421 TTAGTGAATCAAAACAAAGTAATGAACGACTTTTAAATGATACTAATACAGTATCATCTGTA 480
QY 781 AATTCACTCCAAATTTCTCAAAATGGGAAAAATGTTTTCAACACGCAAGATACCTTCAACT 840
Db 481 AATTCACTCCAAATTTCTCAAAATGGGAAAAATGTTTTCAACACGCAAGATACCTTCAACT 540
QY 841 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGATACAGATGCAAGTAATAAGAT 900
Db 541 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGATACAGATGCAAGTAATAAGAT 600
QY 901 GTAGTTAATCAAGCGGTTAATACAAAGTGGCGCTAGAAATGAGAGCATTTAGTTAGCGGCA 960
Db 601 GTAGTTAATCAAGCGGTTAATACAAAGTGGCGCTAGAAATGAGAGCATTTAGTTAGCGGCT 660
QY 961 GTAGCTGCAGATGACCGGAGCTGGCACAGATATTACGAATCAGTTCGCAATGTGACA 1020
Db 661 GTAGCTGCAGATGACCGGAGCTGGCACAGATATTACGAATCAGTTCGCAATGTGAAA 720
QY 1021 GTTGGTATTGACTCTGGTACGACTGTATCCGCAACAGCAGGTTATGTCAAACTGAAT 1080
Db 721 GTTACTATTGACTCTGGTACGACTGTATCCGCAACAGCAGGTTATGTCAAACTGAAT 780
QY 1081 TATGGTTTTTCACTGCTCTAATCTCGTTTAAAGGTGACACATTCAAAATTAATCTGTACCT 1140
Db 781 TATGGTTTTTCACTGCTCTAATCTCGTTTAAAGGTGACACATTCAAAATTAATCTGTACCT 840
QY 1141 AAGAATAATTAACCTTAATGTTGTAACCTCAACTGCTAAAGTGCACCAATTTATGGCTGGA 1200
Db 841 AAGAATAATTAACCTTAATGTTGTAACCTCAACTGCTAAAGTGCACCAATTTATGGCTGGA 900
QY 1201 GATCAAGTATTGCAAAATGTTGTAATCGATAGTGTGTAATGTTATTTATACATTTACA 1260
Db 901 GATCAAGTATTGCAAAATGTTGTAATCGATAGTGTGTAATGTTATTTATACATTTACA 960

QY 1261 GACTATGTAATACTAAAGATGATGTAAGCAACTTTGACCATGCCGCTTATATTGAC 1320
Db 961 GACTATGTTGATCTAAAGAAATGTAACAGCTAATATTAATCTATCCAGCTTATATTGAC 1020
QY 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGGTAGTACAACA 1380
Db 1021 CCTGAAATGTTTAAAGACAGGTAATGTGACATTTGCAACTGGCATAGGAACCAATACT 1080
QY 1381 GCAACAAAAACAGTATTAGTAGATTATGAAAAATATGTTAAGTTTATATAACTTACTATT 1440
Db 1081 GCTAGTAGACAGTATTATCGACTATGAGAAATATGACAATTCCTAATTTATCAATT 1140
QY 1441 AAGAGTCAAAATGACCAATCGATAAAAAACAATAATACGTATCGTTCGTCAGCAAAATTTATGTC 1500
Db 1141 AAGAGTACGATTGATCAAAATCGATAAAAAACAATAATACGTATCGTCGCAAAATTTATGTC 1200
QY 1501 AATCCAAAGTGAGATAAAGTTTATTCGCGCGGTTTTAAACAGGTAATTTAAACCAATACG 1560
Db 1201 AATCCAAAGCGAGATAAAGTTTATTCGCGCGGTTTTAAACAGGTAATTTAAACCAATACG 1260
QY 1561 GATAGTAATGCATTAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 1620
Db 1261 AAGAGTAAAGCTTAATAGATGCAAAAAACACTGATTTAAAGTTTATAGAGTAGATAAT 1320
QY 1621 GCAGCTGATTTATCTGAAAGTTTACTTTGTGTAATCCAGAAAACTTTGAGGATGTCTAAT 1680
Db 1321 GCTAATGATTTATCTGAAAGTTTATTTATGTGAATCTAGCGATTTTGAAGATCTAAT 1380
QY 1681 AGTGTGAATATTACATTTCCCAATCCAAATCAATATAAGTAGATTTAATACGCTGAT 1740
Db 1381 CAAGTGAATTTTCAATTTCCAAATCTAATCAATACAAAGTAGAATTTTCTTACGCGAGAT 1440
QY 1741 GATCAAAATTAACAACCGGTATATAGTAGTTGTGTAATGTCATATTGATCCGAATAGCAAA 1800
Db 1441 GATCAAAATTAACAACCGGTATATTTAGTAGTTGTGTAATGCGCATTTGATCTCTAGACA 1500
QY 1801 GGTGATTTAGCTTTACGTTCAACTTTATATGGGTATAACTGCAATATAATTTGGCGCTCT 1860
Db 1501 GGTGATTTAGCACTACGTTTCGACATTTTATGGTTATGATTTCTAATTTATATGGAGATCT 1560
QY 1861 ATGTCAATGGGACAGAGTAGATTTAATACGATCAGGTTCTGGTGACGGTATCGAT 1920
Db 1561 ATGTCAATGGGACAGAGTAGATTTAATACGATCAGGTTCTGGTGACGGTATCGAT 1620
QY 1921 AAACCAAGTTGTTCTGGAACAACTGATGAGCCTGCTGAAATGAAACCAATTCACAGGAT 1980
Db 1621 AAACCTGTTGTTCTGGAACAACTGATGAGCCTGCTGAAATGAAACCAATTCACAGGAT 1680
QY 1981 TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGATTTCTAATTCAGATAGCGGTTTCAGAT 2040
Db 1681 TCAGATTTCTGACCCAGGTTTCAGATTTCTGCGACGATTTCTAATTCAGATAGCGGTTTCAGAT 1740
QY 2041 TCGGTTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCA 2100
Db 1741 TCGGTTAGTGTATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCG 1800
QY 2101 AGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGATTCGCACTCAGCGAGCGAT 2160
Db 1801 AGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGATTCGCACTCAGCGAGCGAT 1860
QY 2161 TCCGACTCAGCAATGATTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2198
Db 1861 TCAGATTTCCGCGAGTGATTCAGATTCAGCAAGCGATTCAGACTCAGCAAGCGATTCAGAC 1920
QY 2199 ----- 2198
Db 1921 TCAGCAAGCGATTCAGACTCAGACAGTGAATTCAGACAGTGAATTCAGATTCGACAGTGAAC 1980
QY 2199 ----- 2208
Db 1981 TCAGACTCAGATAGCGATTCAGATTTCCGATAGCGATTCGCACTCAGACAGTGAATTCAGAC 2040
QY 2209 AGTGACTCAGATTCGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2268

Db 2041 AGTGATTCAGATTCGACAGCGACTCAGATTCGATAGCGATTCGCACTCAGATAGCGAC 2100
QY 2269 TCAGATTCAGATAGCGATTCAGATTCAGATAGCGATTCAGATTCGACAGTGAATTCGAC 2328
Db 2101 TCAGATTCAGATAGCGATTCGCACTCAGATAGCGACTCAGATTCAGACAGTGAATTCGAC 2160
QY 2329 TCAGACAGCGATTCGCTGACTCCGACAGTGAATTCGCACTCAGACAGCGATTCAGATTCGAC 2388
Db 2161 TCAGATAGCGATTCAGACTCAGACAGTGAATTCGCACTCAGATAGCGATTCAGATTCAGAC 2220
QY 2389 AGTGATTCGCACTCAGATAGCGATTCGCACTCAGATAGCGACTCAGATTCAGACAGCGAT 2448
Db 2221 AGTGACTCGGATTCAGATAGCGATTCGCACTCAGATAGCGATTCAGATTCAGACAGTGAAC 2280
QY 2449 TCAGATTCAGACAGCGATTCAGATTCAGATAGCGATTCAGATTCGCACTCAGACAGTGAATTCAGATTC 2508
Db 2281 TCGGATTCAGATAGCGATTCAGATTCGCACTCAGATAGCGATTCAGATTCAGACAGCGATTCGAC 2340
QY 2509 TCCGACAGTGAATTCGCACTCAGATAGCGATTCAGATTCGCACTCAGATTCGCACTCAGATTCGCACT 2568
Db 2341 TCAGACAGCGATTCGCACTCAGATAGCGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTC 2400
QY 2569 AGTGACTCAGACTCAGACAGTGAATTCGCACTCAGACAGTGAATTCGCACTCAGATTCAGATTCAGATTCAGATTC 2628
Db 2401 AGCGATTCGCACTCAGACAGTGAATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2460
QY 2629 TCCGACTCCGACAGTGAATTCGCACTCAGATAGCGATTCAGATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTC 2688
Db 2461 TCCGACTCAGACAGTGAATTCGCACTCAGATAGCGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2520
QY 2689 TCAGATAGCGATTCGCACTCAGATAGCGATTCAGATTCGCACTCAGACAGCGATTCAGATTCAGATTCAGATTCAGATTC 2748
Db 2521 TCAGATAGCGATTCAGATTCAGATTCAGATTCGCACTCAGACAGCGATTCAGATTCGCACTCAGATTCAGATTCAGATTC 2580
QY 2749 AGCGATTCAGATTCAGACAGCGATTCAGAC----- 2778
Db 2581 AGTGACTCAGATTCAGATAGCGATTCAGATTCAGATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTC 2640
QY 2779 ----- 2790
Db 2641 TCAGATTCAGCGAGTGAATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2700
QY 2791 TCAGATAGTGAATTCGCACTCAGCGAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2850
Db 2701 TCAGGTAGTGAATTCAGATTCAGCGAGTGAATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2760
QY 2851 AGTGATTC----- 2856
Db 2761 AGTGATTCAGACTCAGACAGTGCCTCTGATTCAGCGAGTGAACAGATTCAGACAGCGAC 2820
QY 2857 TCCGACTCAGAAAGTGAATTCAGATTCAGCGATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTC 2916
Db 2821 TCTGACTCAGAAAGTGAATTCAGATTCAGCGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTC 2880
QY 2917 CCGCTAATTCACCTAAATAGTGAATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTC 2976
Db 2881 CCGCTAATTCACCTAAATAGTGAATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTC 2940
QY 2977 AAAGAACCAATTCAGATTCAGATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTC 3036
Db 2941 AAAGAGCCATTCAGATTCAGATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTC 3000
QY 3037 TTATAGCATCAATAGGTTTCAATTCAGATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTC 3096
Db 3001 TTATAGCATCAATAGGTTTCAATTCAGATTCGCACTCAGATTCGCACTCAGATTCGCACTCAGATTCAGATTCAGATTCAGATTCAGATTC 3060
QY 3097 AAA 3099
Db 3061 AAA 3063

ABK13616
 ID ABK13616 standard; DNA; 2792 BP.
 XX
 AC ABK13616;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE S. aureus gene for an antigenic protein #3.
 XX
 KW Antigenic protein; ds; gene; vaccine; SEREX; antibacterial;
 KW antiinflammatory; dermatological; antiulcer; tuberculosis;
 KW immunosuppressive; septicemia; food poisoning; skin disorders;
 KW peritonitis; endocarditis; tuberculosis; blood infection; sepsis;
 KW meningitis; pneumonia; stomach ulcer; gonorrhea; necrotising fasciitis;
 KW impetigo; Lyme's disease; gastro-enteritis; dysentery; shigellosis.
 XX
 OS Staphylococcus aureus
 XX
 XX NC200198499-A1.
 PN
 XX
 PD 27-DEC-2001.
 XX
 PF 20-JUN-2001; 2001WO-GB002685.
 XX
 PR 20-JUN-2000; 2000GB-00014907.
 XX
 XX (UYSH-) UNIV SHEFFIELD.
 PA (BIOS-) BIOSNEXUS INC.
 XX
 XX Foster S, McDowell P, Brummell K, Clarke S;
 XX
 XX WPI; 2002-106544/14.
 XX
 XX Identifying antigenic polypeptides expressed by pathogenic organisms
 PT e.g., Staphylococcus aureus during infection, by SEREX (serological
 PT identification of antigens by recombinant expression cloning) techniques.
 XX
 PS Claim 1; Page 33-34; 85pp; English.
 XX
 CC The invention relates to a method for identifying antigenic polypeptides
 CC expressed by pathogenic organisms e.g., Staphylococcus aureus during
 CC infection, by SEREX (serological identification of antigens by
 CC recombinant expression cloning) techniques. The method involves providing
 CC a nucleic acid library encoding genes/partial gene sequences of
 CC pathogenic organisms, transforming/transfecting the library into host
 CC cells, contacting the polypeptides expressed by the genes with autologous
 CC antisera derived from an animal infected with, or has been infected with
 CC the pathogen and purifying the nucleic acid encoding the polypeptide or
 CC partial polypeptide binding to the antisera. Also included are the
 CC nucleic acids and polypeptides isolated by the method, vectors and
 CC transformed cells expressing them, a vaccine comprising the polypeptide
 CC and the production of monoclonal antibodies against the polypeptides. The
 CC protein and vaccine are useful for immunising an animal (preferably
 CC human) against a pathogenic microbe. The proteins and antibodies are
 CC useful for manufacturing a medicament for treating Staphylococcus aureus-
 CC associated septicemia, food poisoning or skin disorders or
 CC Staphylococcus epidermidis-associated septicemia, peritonitis,
 CC endocarditis, tuberculosis, blood infections, sepsis, meningitis,
 CC pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
 CC Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
 CC sequence is an S. aureus gene isolated by the method of the invention
 CC which encodes an antigenic polypeptide
 XX
 SQ Sequence 2792 BP; 1078 A; 402 C; 525 G; 787 T; 0 U; 0 Other;

Query Match 50.8%; Score 1776.4; DB 6; Length 2792;
 Best Local Similarity 99.9%; Pred. No. 5.5e-311;
 Matches 1788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 GGTACCATTAATACACATCTGCTTTTGAAGGATATGATTTCAAGCTAGGATTACATTA 60
 DB 1003 GGTACCATTAATACACATCTGCTTTTGAAGGATATGATTTCAAGCTAGGATTACATTA 1062

QY 61 GGTAGAGTTCAVATTAAATAATAAAAAATGTTTGCAATCAAATCGTAGCTGTGTTGTA 120
 DB 1063 GGTAGAGTTCAVATTAAATAATAAAAAATGTTTGCAATCAAATCGTAGCTGTGTTGTA 1122
 QY 121 ATTCTTAAATAGCAATAATAATAATGTTTGTAGTAAAGTATATTTGTGGATAATAAAA 180
 DB 1123 ATTCTTAAATAGCAATAATAATAATGTTTGTAGTAAAGTATATTTGTGGATAATAAAA 1182
 QY 181 TATCGATACAAATTAATTCGTATAATGCAATTTTAGTGTATATATTCATTAAACAGAGATT 240
 DB 1183 TATCGATACAAATTAATTCGTATAATGCAATTTTAGTGTATATATTCATTAAACAGAGATT 1242
 QY 241 AAATATATC-TTAAAGGGTATATAGTATAATAATAATGACTTTTAAAAAGAGGGATAAA 299
 DB 1243 AAATATATCTTTAAAGGGTATATAGTATAATAATAATGACTTTTAAAAAGAGGGATAAA 1302
 QY 300 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC 359
 DB 1303 AATGAATATGAAGAAAAAGAAAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTC 1362
 QY 360 AGTGCTTGTAGGTACGTTAATCGGTTTTTGACTACTCAGCAGTAAAGAACAGATGCAAG 419
 DB 1363 AGTGCTTGTAGGTACGTTAATCGGTTTTTGACTACTCAGCAGTAAAGAACAGATGCAAG 1422
 QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAAATGATTCAG 479
 DB 1423 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAAATGATTCAG 1482
 QY 480 TAGCGTTAGTGTGCACCTTAAACACAGACACAAAGTGAGTGATCTAAACATCGTC 539
 DB 1483 TAGCGTTAGTGTGCACCTTAAACACAGACACAAAGTGAGTGATCTAAACATCGTC 1542
 QY 540 ABAACATTAATGCGGAACAGAGTGGCGGCAAAATCCAGACACACAGAAACGACACA 599
 DB 1543 ABAACATTAATGCGGAACAGAGTGGCGGCAAAATCCAGACACACAGAAACGACACA 1602
 QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 659
 DB 1603 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACTAC 1662
 QY 660 GACAAGCAATCAAGCTTAATACACGGCAACAACTCAATCAAGCAATCAAAATCGCGAGGA 719
 DB 1663 GACAAGCAATCAAGCTTAATACACGGCAACAACTCAATCAAGCAATCAAAATCGCGAGGA 1722
 QY 720 ATTAGTGAATCAACAAAGTAAATGAACGACTTTTAAATGATACTAATACAGTATCATCTGT 779
 DB 1723 ATTAGTGAATCAACAAAGTAAATGAACGACTTTTAAATGATACTAATACAGTATCATCTGT 1782
 QY 780 AAATTCACTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGATACTTCAAC 839
 DB 1783 AAATTCACTCAAAATTTCTACAAATGCGGAAATGTTTCAACACGCAAGATACTTCAAC 1842
 QY 840 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATCAAGTATAAAGA 899
 DB 1843 TGAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATCAAGTATAAAGA 1902
 QY 900 TGTAGTTAATCAACGGGTTAATACAGTGGCGCTAGATGAGAGCAATTTAGTTTAGCGGC 959
 DB 1903 TGTAGTTAATCAACGGGTTAATACAGTGGCGCTAGATGAGAGCAATTTAGTTTAGCGGC 1962
 QY 960 AGTAGCTGCAGATGACACGGCAGCTGGCACAGATATTTACGAATCAGTTGACGAATGTGAC 1019
 DB 1963 AGTAGCTGCAGATGACACGGCAGCTGGCACAGATATTTACGAATCAGTTGACGAATGTGAC 2022
 QY 1020 AGTTGGTATTGACTCTGGTACGACTGTGATCCGCCAACAGCGGTTATGTCATAAAGTAA 1079
 DB 2023 AGTTGGTATTGACTCTGGTACGACTGTGATCCGCCAACAGCGGTTATGTCATAAAGTAA 2082
 QY 1080 TTATGGTTTTTTCAGTGGCTAATTTCTGCTGTTAAAGGTGACACATTCAAATTAAGTGTACC 1139
 DB 2083 TTATGGTTTTTTCAGTGGCTAATTTCTGCTGTTAAAGGTGACACATTCAAATTAAGTGTACC 2142
 QY 1140 TAAAGAATTAAACTTAAATGGTGTAACTTCAACTGCTAAAAGTGCCACCAATTTATGGCTGG 1199

Db 2143 TAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGCTAAAGTGCCACCAATTAATGCTGCG 2202
Qy 1200 AGATCAAGTATTGGCAATGGTGAATCGATAGTGTGTAATGTTATTTATACATTAC 1259
Db 2203 AGATCAAGTATTGGCAATGGTGAATCGATAGTGTGTAATGTTATTTATACATTAC 2262
Qy 1260 AGACTATGTAAATACATAAAGATGATGTAAGCAACTTTGACCATGCCCGCTTATATTGA 1319
Db 2263 AGACTATGTAAATACATAAAGATGATGTAAGCAACTTTGACCATGCCCGCTTATATTGA 2322
Qy 1320 CCTGAAATGTTAAAGACAGGTAATGTGATGCTACTGGCATAGGTAAGTACAC 1379
Db 2323 CCTGAAATGTTAAAGACAGGTAATGTGATGCTACTGGCATAGGTAAGTACAC 2382
Qy 1380 AGCAACAAACACAGTATTAGTATGATTAAGAAATATGTAAGTATTTTATACCTTATCTAT 1439
Db 2383 AGCAACAAACACAGTATTAGTATGATTAAGAAATATGTAAGTATTTTATACCTTATCTAT 2442
Qy 1440 TAAAGGTACAAATTGACCAATCGATAAACAATAATACGTATCGTCAGACAAATTTATGT 1499
Db 2443 TAAAGGTACAAATTGACCAATCGATAAACAATAATACGTATCGTCAGACAAATTTATGT 2502
Qy 1500 CAATCAAGTGGAGTAATGTTATGCGCGGTTTAAACAGTAATTTAAACCAATAC 1559
Db 2503 CAATCAAGTGGAGTAATGTTATGCGCGGTTTAAACAGTAATTTAAACCAATAC 2562
Qy 1560 GGATAGTATGATTAATAGATCAGCAAAATCAAGTATTAAGTATTAAGTATAGATAA 1619
Db 2563 GGATAGTATGATTAATAGATCAGCAAAATCAAGTATTAAGTATTAAGTATAGATAA 2622
Qy 1620 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTGAGGATGTCACATAA 1679
Db 2623 TGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAACTTTGAGGATGTCACATAA 2682
Qy 1680 TAGTGTGAATATTACATTCGCAATCCAAATCAATATAAAGTAGTATTAATACGCTGA 1739
Db 2683 TAGTGTGAATATTACATTCGCAATCCAAATCAATATAAAGTAGTATTAATACGCTGA 2742
Qy 1740 TGATCAAAATTAACAACCGTATATAGTATGTTTAAATGTCATATTGATC 1789
Db 2743 TGATCAAAATTAACAACCGTATATAGTATGTTTAAATGTCATATTGATC 2792

RESULT 10

ADA89855

ID ADA89855 standard; DNA; 2792 BP.

AC ADA89855;

XX

XX

DT 20-NOV-2003 (first entry)

XX

DE Staphylococcus aureus antigenic partial gene sequence locus 83.

XX

KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;

KW antibacterial; neuroprotective; immunosuppressive; anti-inflammatory;

KW antitumor; immunostimulant; ophthalmological; pathogenic microbe;

KW bacteraemia; septic shock; organ infection; skin infection;

KW bacterial basal colonisation; bacterial eye infection; septicaemia;

KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;

KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;

KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;

KW gastro-enteritis; dysentery; shigellosis; skin disorder; gene; ds.

XX

OS Staphylococcus aureus.

XX

PN WO2003011899-A2.

XX

PD 13-FEB-2003.

XX

PF 02-AUG-2002; 2002WO-GB003606.

XX

PR 02-AUG-2001; 2001GB-00018825.

PR 09-JAN-2002; 2002GB-00000349.
XX (UYSH-) UNIV SHEPHERD.
PA (BIOS-) BIOSYNEXUS INC.
XX
PI Foster S, Mond J, Clarke S, McDowell P, Brummel K;
XX WPI; 2003-256434/25.
XX
XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
PT useful as a vaccine for immunizing humans against e.g. bacteraemia, septic
PT shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
PT impetigo.
XX
PS Claim 1; Page 104-105; 189pp; English.
XX
CC The present invention describes an antigenic protein or its part, which
CC is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;
CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, neuroprotective, immunosuppressive,
CC antiinflammatory, antitumor, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicaemia, food poisoning,
CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus partial gene
CC sequence from the present invention.
XX
SQ Sequence 2792 BP; 1078 A; 402 C; 525 G; 787 T; 0 U; 0 Other;

Query Match 50.8%; Score 1776.4; DB 8; Length 2792;
Best Local Similarity 99.9%; Pred. NO. 5.5e-311;
Matches 1788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GGTACCAATAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGATTACATTA 60
Db 1003 GGTACCAATAATTACACATCTGCTTTTGAAAAAATATGATTTCAAGCTAGATTACATTA 1062
Qy 61 GGTAGAGTTCATATTATAATAAAAAATGTTTGCATCAAAATCGTACGTTTCGTTGTA 120
Db 1063 GGTAGAGTTCATATTATAATAAAAAATGTTTGCATCAAAATCGTACGTTTCGTTGTA 1122
Qy 121 ATTCCTAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTATTTGGATATAAAA 180
Db 1123 ATTCCTAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTTATTTGGATATAAAA 1182
Qy 181 TATCGATCAAAATTAATTTGCTATATGCAATTTAGTGTATATAATTCATTACAGAGATT 240
Db 1183 TATCGATCAAAATTAATTTGCTATATGCAATTTAGTGTATATAATTCATTACAGAGATT 1242
Qy 241 AAAATATATC-TTAAAGGTATATAGTTAATAATAAAATGACTTTTTAAAAAGAGGAATAA 299
Db 1243 AAAATATATCTTTAAAGGGTATATAGTTAATAATAAAATGACTTTTTAAAAAGAGGAATAA 1302
Qy 300 AATGAATATGAGAAAAGAAAACACGCAATTCGGAAAAAATCGATTGGCGGCTTC 359
Db 1303 AATGAATATGAGAAAAGAAAACACGCAATTCGGAAAAAATCGATTGGCGGCTTC 1362

```
QY 360 AGTCTGTAGGTACGTTAAATCGGTTTTCGACTACTCAGCAGTAAAGACGATGCAAG 419
Db 1363 AGTCTGTAGGTACGTTAAATCGGTTTTCGACTACTCAGCAGTAAAGACGATGCAAG 1422
QY 420 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCAAG 479
Db 1423 TGAATAATAGTGTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCAAG 1482
QY 480 TAGCGTTAGTGTGACCTTAAACACAGACGACAAACGTTGAGTGATCTAATAACATCGTC 539
Db 1483 TAGCGTTAGTGTGACCTTAAACACAGACGACAAACGTTGAGTGATCTAATAACATCGTC 1542
QY 540 AAACACTAATAATGCGGAACGAGTGTGGCGCAAAATCCAGACCAACAGGAAACGACACA 599
Db 1543 AAACACTAATAATGCGGAACGAGTGTGGCGCAAAATCCAGACCAACAGGAAACGACACA 1602
QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGGTGAAGCTACTACTAC 659
Db 1603 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGGTGAAGCTACTACTAC 1662
QY 660 GACAAAGATCAAGCTAATAACACGCGCAACAACTCAATCAAGCAATACAAATCGGAGGA 719
Db 1663 GACAAAGATCAAGCTAATAACACGCGCAACAACTCAATCAAGCAATACAAATCGGAGGA 1722
QY 720 ATTAGTGAATCAAAACAGTAATGAACGACTTTTATGATACTAATAACAGTATCATCTGT 779
Db 1723 ATTAGTGAATCAAAACAGTAATGAACGACTTTTATGATACTAATAACAGTATCATCTGT 1782
QY 780 AAATTCACCTCAAAATTTCTCAAAATCGGAAATGTTTCAACACGCAAGATACTTCAAC 839
Db 1783 AAATTCACCTCAAAATTTCTCAAAATCGGAAATGTTTCAACACGCAAGATACTTCAAC 1842
QY 840 TGAAGCAACACCTTCAAAACAGTAATGAACGACTCCACAGAGTACAGATCAAGTAATAAGA 899
Db 1843 TGAAGCAACACCTTCAAAACAGTAATGAACGACTCCACAGAGTACAGATCAAGTAATAAGA 1902
QY 900 TGTAGTTAATCAAGCGGTTAATCAAGTGGCGCTAGAAATCAGAGACTTTAGTTTAGCGGC 959
Db 1903 TGTAGTTAATCAAGCGGTTAATCAAGTGGCGCTAGAAATCAGAGACTTTAGTTTAGCGGC 1962
QY 960 AGTAGCTGCAGATGCACGGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGAC 1019
Db 1963 AGTAGCTGCAGATGCACGGCAGCTGGCACAGATATTACGAATCAGTTGACGAATGTGAC 2022
QY 1020 AGTTGGTATTGACTCTGGTAGCAGTGTGATCCGACCAGCAGGTTATGTCACAACTGAA 1079
Db 2023 AGTTGGTATTGACTCTGGTAGCAGTGTGATCCGACCAGCAGGTTATGTCACAACTGAA 2082
QY 1080 TTATGGTTTTTTCAGTGGCTAATTTCTGCTGTTAAAGGTGACACATTCACAAATTAAGTACC 1139
Db 2083 TTATGGTTTTTTCAGTGGCTAATTTCTGCTGTTAAAGGTGACACATTCACAAATTAAGTACC 2142
QY 1140 TAAAGAATTAACCTTAAATGGTGTAACTTCAACTGTCTAAAGTGCACCAATTAAGTGTG 1199
Db 2143 TAAAGAATTAACCTTAAATGGTGTAACTTCAACTGTCTAAAGTGCACCAATTAAGTGTG 2202
QY 1200 AGATCAAGTATTGCAATGGTGTAAATCGATAGTGGTAAGTTATTTATACATTTAC 1259
Db 2203 AGATCAAGTATTGCAATGGTGTAAATCGATAGTGGTAAGTTATTTATACATTTAC 2262
QY 1260 AGACTATGTAATAACTAAAGATGATGAAAGCAACCTTTCACCATGCCGCTTATATTGA 1319
Db 2263 AGACTATGTAATAACTAAAGATGATGAAAGCAACCTTTCACCATGCCGCTTATATTGA 2322
QY 1320 CCCTGAAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGTAGTACAAC 1379
Db 2323 CCCTGAAAATGTTAAAAAGACAGGTAATGTGACATTTGGCTACTGGCATAGTAGTACAAC 2382
QY 1380 AGCAACAAACAGTATTAGTAGATTATGAAAAATATGGAAGTTTATTAACCTTATCTAT 1439
Db 2383 AGCAACAAACAGTATTAGTAGATTATGAAAAATATGGAAGTTTATTAACCTTATCTAT 2442
```

```
QY 1440 TAAAGGTACAATTGACCAAAATCGATAAAACAAATAATAACGTATCGTCAGACAAATTTATGT 1499
Db 2443 TAAAGGTACAATTGACCAAAATCGATAAAACAAATAATAACGTATCGTCAGACAAATTTATGT 2502
QY 1500 CAATCCAAGTGGAGATAAAGCTTTATTCGCGCGGTTTTTAAACAGGTAATTTAAAAACCAATAC 1559
Db 2503 CAATCCAAGTGGAGATAAAGCTTTATTCGCGCGGTTTTTAAACAGGTAATTTAAAAACCAATAC 2562
QY 1560 GGNATGTAATGCTAATTAAGATCAGCAAAATACAAAGTATTAAGTATATAAAGTAGATAA 1619
Db 2563 GGNATGTAATGCTAATTAAGATCAGCAAAATACAAAGTATTAAGTATATAAAGTAGATAA 2622
QY 1620 TGCAGCTGATTTTATCTGAAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCATAA 1679
Db 2623 TGCAGCTGATTTTATCTGAAAAGTTACTTTGTGAATCCAGAAAACCTTTGAGGATGTCATAA 2682
QY 1680 TAGTGTGAATATTACATTTCCAAAATCCAAATCAATATAAAGTAGAGTTTAATAGCGCTGA 1739
Db 2683 TAGTGTGAATATTACATTTCCAAAATCCAAATCAATATAAAGTAGAGTTTAATAGCGCTGA 2742
QY 1740 TGATCAAAATTACAAACACCGGTATATAGTAGTTGTTAATGGTCATATTGATC 1789
Db 2743 TGATCAAAATTACAAACACCGGTATATAGTAGTTGTTAATGGTCATATTGATC 2792

RESULT 11
AAD46861
ID AAD46861 standard; DNA; 1560 BP.
XX
AC AAD46861;
XX
DT 27-JAN-2003 (first entry)
XX
DE Staphylococcus aureus Clf40 DNA.
XX
KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy;
KW gene; ds
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..1560
FT /tag= a
FT /product= "Staphylococcus aureus Clf40 protein"
FT /note= "CDS does not include start and stop codon"
XX
XX WO200272600-A2.
XX
XX 19-SEP-2002.
XX
XX 28-JAN-2002; 2002WO-US002296.
XX
XX 26-JAN-2001; 2001US-0264072P.
XX
XX 12-MAR-2001; 2001US-027461P.
XX
XX 18-JUN-2001; 2001US-029843P.
XX
XX 30-JUL-2001; 2001US-0308416P.
XX
XX (INHI-) INHIBITEX INC.
XX
XX Patti-JM Hutchins JT, Domanski P, Patel P, Hall A;
XX
XX WPI; 2002-759834/82.
XX
XX P-PSDB; AAE29262.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
XX treating or preventing Staphylococcus aureus infection e.g. wound
XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
XX a human or animal.
XX
XX Claim 10; Page 66-67; 80pp; English.
PS
```


XX The invention relates to monoclonal antibody which binds the clumping
CC factor A (ClfA) protein from *Staphylococcus aureus*. The anti-ClfA
CC monoclonal antibody is useful for treating or preventing *S. aureus*
CC infection in a human or animal, and for inhibiting the binding of
CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
CC *S. aureus* Clf40 protein, *S. aureus* Clf33 protein, or the *S. aureus* N3
CC protein is useful for inducing an immunological response in a human or
CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is *Staphylococcus aureus* Clf40 DNA
XX
SQ Sequence 1560 BP; 584 A; 269 C; 299 G; 408 T; 0 U; 0 Other;
Query Match 44.6%; Score 1560; DB 6; Length 1560;
Best Local Similarity 100.0%; Pred. No. 5.5e-272; Mismatches 0; Gaps 0;
Matches 1560; Conservative 0; Indels 0;
418 AGTGAATAAGTGTTCACCAATCTCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCA 477
Db 1 AGTGAATAAGTGTTCACCAATCTCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATCA 60
478 AGTAGCGTGTAGTGTGACCTTAACACAGACGACACAAAGCTGAGTGATCTAAACATCG 537
Db 61 AGTAGCGTGTAGTGTGACCTTAACACAGACGACACAAAGCTGAGTGATCTAAACATCG 120
538 TCAACACTAATAATGGGAAACGAGTGTGGCGCAAAATCCAGCACACACGAAACGACA 597
Db 121 TCAACACTAATAATGGGAAACGAGTGTGGCGCAAAATCCAGCACACACGAAACGACA 180
598 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACT 657
Db 181 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGGTGAAGCTACTACT 240
658 ACGCAACGAATCAGCTAATACACCGCGCAACACTCAATCAGCAATCAAAATCGCGAG 717
Db 241 ACGCAACGAATCAGCTAATACACCGCGCAACACTCAATCAGCAATCAAAATCGCGAG 300
718 GAATAGTGAATCAACAAAGTAATCAACAGCACTTTTAATGATCTAATACAGTATCATCT 777
Db 301 GAATAGTGAATCAACAAAGTAATCAACAGCACTTTTAATGATCTAATACAGTATCATCT 360
778 GTAAATTCACCTCAAAATCTCAAAATCGGAATGTTTCAACCAACCAAGATACCTCA 837
Db 361 GTAAATTCACCTCAAAATCTCAAAATCGGAATGTTTCAACCAACCAAGATACCTCA 420
838 ACTGAGCAACACCTTCAACCAATGAATCAGTCCACAGGTACAGATGCAAGTAATAA 897
Db 421 ACTGAGCAACACCTTCAACCAATGAATCAGTCCACAGGTACAGATGCAAGTAATAA 480
898 GATGTAGTGAATCAACAGCGTTAATAACAAGTGGCGCTAGAAATCAGAGCATTTAGTTAGCG 957
Db 481 GATGTAGTGAATCAACAGCGTTAATAACAAGTGGCGCTAGAAATCAGAGCATTTAGTTAGCG 540
958 GCAGTAGCTGCAGATGCAACCGGACAGTGGCAAGATATACGAATCAGTTCAGCAATG 1017
Db 541 GCAGTAGCTGCAGATGCAACCGGACAGTGGCAAGATATACGAATCAGTTCAGCAATG 600
1018 ACAGTTGGTATGACTCTGTGACGATCTGATCGCACCAAGCAGGTTATGCAACTG 1077
Db 601 ACAGTTGGTATGACTCTGTGACGATCTGATCGCACCAAGCAGGTTATGCAACTG 660
1078 AATTATGTTTTTTCAGTGCCTAAATCTGCTGTAAAGGTGACACATTCACAAATACCTGTA 1137
Db 661 AATTATGTTTTTTCAGTGCCTAAATCTGCTGTAAAGGTGACACATTCACAAATACCTGTA 720
1138 CCTAAGAAATTAACCTTAATGGTGTAACTTCACTGCTAAAGTGCACCAATATGGCT 1197
Db 721 CCTAAGAAATTAACCTTAATGGTGTAACTTCACTGCTAAAGTGCACCAATATGGCT 780
1198 GGAGATCAAGTATGGCAAAATGGTGAATCATGATGATGTAATGTTATTTATACATTT 1257
Db 781 GGAGATCAAGTATGGCAAAATGGTGAATCATGATGATGTAATGTTATTTATACATTT 840

QY 1258 ACAGACTATGTAATAATACATAAGATGATGTAAAGCAACTTTTGACCATGCCCGCTATATT 1317
Db 841 ACAGACTATGTAATAATACATAAGATGATGTAAAGCAACTTTTGACCATGCCCGCTATATT 900
QY 1318 GACCCCTGAAAATGTTTAAAAAGACAGGTAATGTGACATTGGCTACTGGCATAGGTAGTACA 1377
Db 901 GACCCCTGAAAATGTTTAAAAAGACAGGTAATGTGACATTGGCTACTGGCATAGGTAGTACA 960
QY 1378 ACAGCAAAACAAACAGTATTTAGTAGATATGAAAAATATGGTAAGTATTAACCTATCT 1437
Db 961 ACAGCAAAACAAACAGTATTTAGTAGATATGAAAAATATGGTAAGTATTAACCTATCT 1020
QY 1438 ATTAAGGTACAAATGACCAATCGATAAACAATAATACGATCGTCGACACAAATTTAT 1497
Db 1021 ATTAAGGTACAAATGACCAATCGATAAACAATAATACGATCGTCGACACAAATTTAT 1080
QY 1498 GTCAATCAAGTGGAGATAACGTTATTTCGCGGGTTTTAACAGGTAATTTAAAAACCAAT 1557
Db 1081 GTCAATCAAGTGGAGATAACGTTATTTCGCGGGTTTTAACAGGTAATTTAAAAACCAAT 1140
QY 1558 ACGGATAGTATGCAATTAATAGATCAGCAAAATACAGTATTAAGTATTAAGTAGAT 1617
Db 1141 ACGGATAGTATGCAATTAATAGATCAGCAAAATACAGTATTAAGTATTAAGTAGAT 1200
QY 1618 AATGAGCTGATTTATCTGAAAGTTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCAT 1677
Db 1201 AATGAGCTGATTTATCTGAAAGTTTACTTTGTGAATCCAGAAACCTTTGAGGATGTCAT 1260
QY 1678 AATAGTGTGAATTTTACATTCCTCAATCCAAATCAATATAAGTAGAGTATTAATACGCT 1737
Db 1261 AATAGTGTGAATTTTACATTCCTCAATCCAAATCAATATAAGTAGAGTATTAATACGCT 1320
QY 1738 GATGATCAAAATACAGACCGGTATATAGTAGTGTGTTATGTCATATTGATCCGATAGC 1797
Db 1321 GATGATCAAAATACAGACCGGTATATAGTAGTGTGTTATGTCATATTGATCCGATAGC 1380
QY 1798 AAAGTGATTTAGCTTTTCACTTTATATGGGTATACTCGAATATAATTTGGCGC 1857
Db 1381 AAAGTGATTTAGCTTTTCACTTTATATGGGTATACTCGAATATAATTTGGCGC 1440
QY 1858 TCTATGTCATGGGACAAACGAAAGTAGCATTTAATAACGATCAGGTTCTGGTGACGGTATC 1917
Db 1441 TCTATGTCATGGGACAAACGAAAGTAGCATTTAATAACGATCAGGTTCTGGTGACGGTATC 1500
QY 1918 GATAAACAGTGTCTTCCGACCAACCTGATGAGCTGTGAAATTTGAACCAATTTCCAGAG 1977
Db 1501 GATAAACAGTGTCTTCCGACCAACCTGATGAGCTGTGAAATTTGAACCAATTTCCAGAG 1560

RESULT 12
AAD46862
ID AAD46862 standard; DNA; 990 BP.
XX
XX AAD46862;
XX
XX 27-JAN-2003 (first entry)
XX
DE *Staphylococcus aureus* Clf33 DNA.
XX
XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;
KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy;
XX gene; ds.
XX
OS *Staphylococcus aureus*.
XX
XX Key Location/Qualifiers
PH 1. .990
FT /*tag= a
FT /product= "Staphylococcus aureus Clf33 protein"
FT /transl_except= (pos:1. .3, aa:Met-Val)
FT /note= "CDS does not include start and stop codon"

FT XX DN WO200272600-A2. /partial
XX XX PD 19-SEP-2002.
XX XX PF 28-JAN-2002; 2002WO-US002296.
XX XX PR 26-JAN-2001; 2001US-0264072P.
XX XX PR 12-MAR-2001; 2001US-0274611P.
XX XX PR 18-JUN-2001; 2001US-0298413P.
XX XX PR 30-JUL-2001; 2001US-0308116P.
XX XX PA (INH1-) INHIBITEX INC.
XX XX PI Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
XX XX DR WPI: 2002-759834/82.
XX XX DR P-PSDB: AAE29263.
XX XX PT New anti clumping factor A (ClfA) monoclonal antibody, useful for
XX XX PT treating or preventing Staphylococcus aureus infection e.g. wound
XX XX PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
XX XX PT a human or animal.
XX XX PS Claim 10; Page 70; 80pp; English.
XX XX CC The invention relates to monoclonal antibody which binds the clumping
XX XX CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
XX XX CC monoclonal antibody is useful for treating or preventing S. aureus
XX XX CC infection in a human or animal, and for inhibiting the binding of
XX XX CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
XX XX CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
XX XX CC protein is useful for inducing an immunological response in a human or
XX XX CC animal. These staphylococcal infections include wound infections, sepsis,
XX XX CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
XX XX CC present sequence is Staphylococcus aureus Clf33 DNA
XX XX SQ Sequence 990 BP; 352 A; 153 C; 192 G; 293 T; 0 U; 0 Other;
Query Match 28.3%; Score 990; DB 6; Length 990;
Best Local Similarity 100.0%; Pred. No. 3.4e-169;
Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 961 GTAGCTCAGATGACACCGGACGCTGGACAGATATTACGATGATGACGAAATGTGACA 1020
DB 1 GTAGCTCAGATGACACCGGACGCTGGACAGATATTACGATGATGACGAAATGTGACA 60
QY 1021 GTTGGTATTGACTGTGCTGACGCTGTATCCGACCAAGCAGGTTATGTCAAACCTGAAT 1080
DB 61 GTTGGTATTGACTGTGCTGACGCTGTATCCGACCAAGCAGGTTATGTCAAACCTGAAT 120
QY 1081 TATGGTTTTTCAGTGCCTTAATTTCTGCTGTTAAAGGTGACACATTCAAATAAAGTGTACCT 1140
DB 121 TATGGTTTTTCAGTGCCTTAATTTCTGCTGTTAAAGGTGACACATTCAAATAAAGTGTACCT 180
QY 1141 AAGAATAAATTAATTAATGTTGTAATCTCACTGCTAAAGTGCACCAATTAATGCTGGA 1200
DB 181 AAGAATAAATTAATTAATGTTGTAATCTCACTGCTAAAGTGCACCAATTAATGCTGGA 240
QY 1201 GATCAAGTATTGGCAATGGTGTATCGATAGTGTATGTTATTTATACATTTACA 1260
DB 241 GATCAAGTATTGGCAATGGTGTATCGATAGTGTATGTTATTTATACATTTACA 300
QY 1261 GACTATGTAATATCAAGATGATGTAAGAACCACTTTGACCATGCCGCTTATATTGAC 1320
DB 301 GACTATGTAATATCAAGATGATGTAAGAACCACTTTGACCATGCCGCTTATATTGAC 360
QY 1321 CCTGAAATGTTAAAGACAGGTAATGTGATTTGGCTAGTGGATAGTACACA 1380
DB 361 CCTGAAATGTTAAAGACAGGTAATGTGATTTGGCTAGTGGATAGTACACA 420
QY 1381 GCAACAAACAGTATTAGTAGATTATGAAAAATATGTTAAAGTATTATACCTTATT 1440

DB 421 GCAACAAACAGTATTAGTAGATTATGAAAAATATGTTAAAGTATTATACCTTATT 480
QY 1441 AAAGGTACAATTGACCAATCGATAAAACAAATAATACGTATCGTCAGACAAATTATGTC 1500
DB 481 AAAGGTACAATTGACCAATCGATAAAACAAATAATACGTATCGTCAGACAAATTATGTC 540
QY 1501 AATCCAAGTGGAGATAACGTTATTGGCCCGGTTTTAACAGGTAATTTAAAAACCAATAACG 1560
DB 541 AATCCAAGTGGAGATAACGTTATTGGCCCGGTTTTAACAGGTAATTTAAAAACCAATAACG 600
QY 1561 GATAGTAATGCATTAAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 1620
DB 601 GATAGTAATGCATTAAATAGATCAGCAAAATACAAGTATTAAAGTATATAAAGTAGATAAT 660
QY 1621 GCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTGGAGGATGTCACCTAAT 1680
DB 661 GCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAAACCTTGGAGGATGTCACCTAAT 720
QY 1681 AGTGTGAATATTACATTCCTCCCAATCCAAATCAATATAAAGTAGAGTTTAATACGCTGAT 1740
DB 721 AGTGTGAATATTACATTCCTCCCAATCCAAATCAATATAAAGTAGAGTTTAATACGCTGAT 780
QY 1741 GATCAAAATTACAACCGGTATATAGTAGTTGTTAAAGGTCAATATGATCCGAATAGACAA 1800
DB 781 GATCAAAATTACAACCGGTATATAGTAGTTGTTAAAGGTCAATATGATCCGAATAGACAA 840
QY 1801 GGTGATTTAGCTTTACGTTCAACTTTATATATGGGTATTAACCTGATATAATTTGGCGCTCT 1860
DB 841 GGTGATTTAGCTTTACGTTCAACTTTATATATGGGTATTAACCTGATATAATTTGGCGCTCT 900
QY 1861 ATGTCATGGGACAAACGAAGTAGCATTTAATAACGATCAGGTTCTGCTGACGGTATCGAT 1920
DB 901 ATGTCATGGGACAAACGAAGTAGCATTTAATAACGATCAGGTTCTGCTGACGGTATCGAT 960
QY 1921 AAACCAAGTTCTTCTGTAACCAACCTGATGAG 1950
DB 961 AAACCAAGTTCTTCTGTAACCAACCTGATGAG 990
RESULT 13
AAZ93533
ID AAZ93533 standard; DNA; 5406 BP.
XX AC AAZ93533;
XX DT 24-JUL-2000 (first entry)
XX DE Cell wall protein SdrF coding sequence.
XX KW SdrF; SdrH; SdrI; coagulase negative; staphylococcus; septicemia;
XX KW osteomyelitis; endocarditis; immune response; vaccine; graft; stent;
XX KW intravenous catheter; heart valve; cardiac; ss.
XX OS Staphylococcus sp.
PH Key Location/Qualifiers
FT CDS 1..5406
FT FT /*tag= a
FT FT /label= SdrF cell wall protein
FT FT 34..36
FT FT /*tag= b
FT FT /transl_except= TAA stop codon
FT FT 82..84
FT FT /*tag= c
FT FT /transl_except= TAA stop codon
FT FT 5311..5313
FT FT /*tag= d
FT FT /transl_except= TAG stop codon
FT FT 5320..5322
FT FT /*tag= e
FT FT /transl_except= TGA stop codon
FT FT 5377..5379

```
FT      /*tag= f
FT      /transl_except= TGA stop codon
FT      5398..5400
FT      /*tag= g
FT      /transl_except= TGA stop codon
XX
XX      WO200012689-A1.
XX      09-MAR-2000.
XX
XX      31-AUG-1999; 99WO-US019728.
XX
XX      31-AUG-1998; 98US-0098443P.
XX      25-JAN-1999; 99US-0117119P.
XX
XX      (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX      (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
XX      Poster TJ, Hook M, Davis S, Hartford O, McCrear K, Ni Eidhin D;
XX      P-PSDB; AAY83170.
XX
XX      WPI; 2000-256637/22.
XX      DR
XX      P-PSDB; AAY83170.
XX
XX      Recombinant or synthetic proteins from coagulase-negative staphylococci
XX      useful for prevention, treatment and diagnosis of staphylococcal
XX      infections bind soluble and immobilized fibrinogen.
XX
XX      Claim 3; Fig 2; 104pp; English.
XX
XX      Isolated staphylococcus-sdr cell wall proteins which bind both soluble
XX      and immobilized fibrinogen are useful for treating or preventing
XX      coagulase-negative staphylococcal infection such as septicemia,
XX      osteomyelitis or endocarditis, and for inducing immune responses in
XX      patients. The cell wall proteins are also useful for reducing coagulase-
XX      negative staphylococci infection of indwelling medical devices such as
XX      vascular grafts, vascular stents, intravenous catheters, artificial heart
XX      valves and cardiac assist devices. The cell wall associated proteins are
XX      able to inhibit staphylococcal adhesion to immobilised extracellular
XX      matrix or host cells present on the surface of implanted biomaterials
XX
XX      Sequence 5406 BP; 1982 A; 955 C; 1138 G; 1331 T; 0 U; 0 Other;
XX
XX      Query Match 17.7%; Score 619.2; DB 3; Length 5406;
XX      Best Local Similarity 78.8%; Pred. No. 3.5e-102;
XX      Matches 738; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
XX
XX      1978 GATTTCAGATCTGACCCAGGTTTCAGATTCGCGAGGATTCCTAATTCAGATAGCGGTTC 2037
XX      4246 GACTCAGACGAGATAGTATTCGATTCGCGATTCGCGATTCGATTCGATTCGATTCGATTC 4305
XX
XX      2038 GATTCGGGTAGTGTATTCATCATCAGATAGTGTGTTTCAGATTCAGCGAGTGTTCAGATTC 2097
XX      4306 GACTCAGACAGTACTCAGATTCGATAGCGACTCGGATTCAGATAGTGTTCGACGCA 4365
XX
XX      2098 GCAAGTGTATTCAGACTCAGCGAGTGTTCAGATTCAGCAAGCGATTCGAGTTCAGCGAGC 2157
XX      4366 GACAGTGTATTCAGACTCAGATAGTGTTCGATTCGCGATTCGATTCGCGAGCAGCAGC 4425
XX
XX      2158 GATTCGCGACTCAGCAAGTGTTCGATTCAGATAGCGATTCGATTCAGATAGCGATTCAGTCA 2217
XX      4426 GATTCGACTCAGATAGTGTTCGATTCAGACGAGATAGTGTTCGATTCGATTCGATTCGATTCG 4485
XX
XX      2218 GATTCGCGAGTGTATTCAGATTCAGATAGCGATTCGATTCAGACAGTGTTCAGATTCAGTCA 2277
XX      4486 GATGCAGACAGCGACTCGGATTCAGATAGTGTTCGATTCGCGAGCAGCAGTGTTCAGACTCA 4545
XX
XX      2278 GATACCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGTTCGATTCGAGTTCAGACAGC 2337
XX      4546 GATAGTGTATTCGATTCGATAGTGTTCGATTCGCGAGCAGCAGTGTTCGATTCGATTCGATTCG 4605
XX
XX      2338 GATTCGACTCAGCGAGTGTTCGATTCGAGTTCAGATTCGAGCAGCGATTCAGATTCGATTCGATTC 2397
XX      4606 GATTCAGACTCAGCGAGCGATTCAGATTCAGACAGCGAGTTCAGATTCGATTCGATTCGATTC 4665
```

```
QY      2398 GACTCAGATAGCGATTCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTC 2457
DB      4666 GATTTCAGACAGTGTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 4725
QY      2458 GACAGCGATTCAGATTCAGATAGCGATTCAGATTCGCGAGTGTTCAGATTCGCGAGT 2517
DB      4726 GATAGCGACTCAGATTCAGACAGTGTTCGCGACTCAGATAGTGTTCGCGATTCGAGT 4785
QY      2518 GACTCGGATTCAGATAGCGATTCAGATTCGCGAGTGTTCAGATTCGCGAGTGTTCAG 2577
DB      4786 GATTTCGATTCGATAGCGATTCGATTCGCGATTCGATTCGCGATTCGAGTGTTCG 4845
QY      2578 GACTCAGACAGTGTTCGATTCGCGAGTGTTCAGATTCGCGATTCAGATAGTGTTCG 2637
DB      4846 GACTCAGACAGCGACTTCGATTCAGATAGTGTTCGACTCAGACAGCGATTCGAGTTC 4905
QY      2638 GACAGTGTTCGATTCAGATAGCGACTCAGATTCGCGAGTGTTCGATTCGCGATTCAG 2697
DB      4906 GATAGTGTTCGATTCAGACAGTGTTCGCGACTCAGACAGCGACTTCGCGATTCAGAT 4965
QY      2698 GATTTCGCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAG 2757
DB      4966 GATTTCGCGAGCGAGCGACTTCGATTCAGATAGTGTTCGCGAGCGAGCGATTCG 5025
QY      2758 GATTTCAGACAGCGACTCAGACAGTGTTCAGATTCAGATAGTGTTCGCGATTCAG 2817
DB      5026 GATAGTGTTCGATTCAGACAGTGTTCGCGACTCAGACAGCGATTCGAGTTCAGACAG 5085
QY      2818 GATTTCAGACTCAGGATTCGATTCGCGATTCAGATTCAGATTCGAGTGTTCGAGTGTTC 2877
DB      5086 GACTTCGCGACTCAGATAGCGACTTCGCGATTCAGACAGTGTTCGCGACTCAGATAGT 5145
QY      2878 AATAGCGATTCGCGACTCAGGTTCTTAACATATGTA 2913
DB      5146 GATTTCAGACAGCGACTTCGATTCGATAGTGTTCGATAGTGTTCGATAGTGTTCGATTC 5181

RESULT 14
AAZ51201
ID AAZ51201 standard; DNA; 5406 BP.
XX
AC AAZ51201;
XX
DT 06-JUN-2000 (first entry)
XX
DE Staph. epidermidis serine-aspartate repeat region protein SdrF gene.
KW Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;
KW microbial surface components recognising adhesive matrix molecules;
KW collagen binding protein; CBP; CNA; fibrinogen binding protein;
KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;
KW fibronectin binding protein; Staphylococcus infection;
KW serine-aspartate repeat region protein; SDR protein; SdrF; ds.
XX
OS Staphylococcus epidermidis.
XX
FH Key Location/Qualifiers
CDS 1..5406
FT /*tag= a
FT /product= "SdrF protein"
FT /transl_except= (pos:34..36, aa:Xaa)
FT /transl_except= (pos:82..84, aa:Xaa)
FT /transl_except= (pos:5311..5313, aa:Xaa)
FT /transl_except= (pos:5320..5322, aa:Xaa)
FT /transl_except= (pos:5377..5379, aa:Xaa)
FT /transl_except= (pos:5398..5400, aa:Xaa)
FT /note= "The coding region does not include stop codon.
FT xaa corresponds to in-frame stop codon"
FT /partial
XX
XX      WO200012131-A1.
```

PD 09-MAR-2000.
XX
PF 31-AUG-1999; 99WO-US019727.
XX
XX 31-AUG-1998; 98US-0098439P.
XX
XX (INH-) INHIBITEX INC.
PA (TEXA) UNIV TEXAS A & M SYSTEM
PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
XX Patti JM, Foster TJ, Hook M;
PI WPI; 2000-237781/20.
XX P-PSDB; AAY70119.
XX
XX Composition used for generating immune response or for inhibiting
PT microbial colonization in an animal comprises antibodies that bind
PT collagen binding protein, fibrinogen binding protein and, optionally,
PT fibronectin binding protein.
XX
XX Disclosure; Fig 3; 115pp; English.
XX
XX The patent discloses multicomponent vaccines containing selected
CC combinations of bacterial binding proteins termed MSCRAMM (microbial
CC surface components recognising adhesive matrix molecules) or their
CC antibodies. A vaccine composition is provided that includes collagen
CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein
CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB), and
CC optionally a fibronectin binding protein e.g. FnBP-A. The vaccines are
CC useful for imparting protection against a broad spectrum of
CC Staphylococcal strains and for inhibiting microbial colonisation,
CC especially of Staphylococcus aureus, in an animal. The combinations can
CC also be used to select donor blood pools for the preparation of purified
CC blood products for passive immunisation. The present sequence is a gene
CC encoding serine-aspartate repeat region protein, SdrF from Staphylococcus
CC epidermidis. The Sdr protein is useful in vaccine preparation in
CC combination with specific bacterial binding proteins. These vaccines can
CC be used to treat a broad spectrum of bacterial infections, including
CC those arising from both coagulase-positive and coagulase-negative
CC bacteria
XX
SQ Sequence 5406 BP; 1982 A; 955 C; 1138 G; 1331 T; 0 U; 0 Other;

Query Match 17.7%; Score 619.2; DB 3; Length 5406;
Best Local Similarity 78.8%; Pred. No. 3.5e-102;
Matches 738; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 1978 GATTTCAGATTCTGACCCAGGTTTCAGATTCTGGCAGCGGATTCTAATTCAGATAGCGGTTCA 2037
DB 4246 GACTCAGACGACATAGTGTGATTCGGATTCCGATAGCGGATTTCAGATTCTGATAGTACTCA 4305
QY 2038 GATTCGGGTAGTGTATTCATCAGATAGTAGTGGTTTCAGATTTCAGGAGTGTATTCAGATTCA 2097
DB 4306 GACTCAGACAGTACTCAGATTCCGATAGCGACTCCGATTTCAGATAGTGTATTCGAGCGCA 4365
QY 2098 GCAGGTGATTTCAGACTCAGCGAGTGTATTCAGATTTCAGCAGCGGATTCCGACTCAGCGAGC 2157
DB 4366 GACAGTACTCAGACTCAGATAGTAGTCTCGATTCCGATAGTGTATTCGAGCGGACGACG 4425
QY 2158 GATTCGAGCTCAGACATGACTCGGATTTCAGATAGCGGATTCTGACTCAGACAGTGTACTCA 2217
DB 4426 GATTCGACTCAGATAGTACTCAGACGCGAGATAGTGTATTCGATTCCGATAGCGGATTTCG 4485
QY 2218 GATTCGACAGTGTACTCAGATTTCAGATAGCGGATTCTGACTCAGACAGTGTACTCAGATTCA 2277
DB 4486 GATTCGACAGCGACTCCGATTTCAGATAGTGTATTCGAGCGGATTTCGAGTGTACTCAGACTCA 4545
QY 2278 GATGAGGATTTCAGATTTCAGATAGCGGATTTCGAGTGTATTCGAGTGTACTCAGACAGC 2337
DB 4546 GATAGTACTCCGATTTCGATAGTGTATTCGAGCGGATTTCGAGTGTACTCAGATTTCGATAGC 4605
QY 2338 GATTCTGACTCCGACAGTGTATTCGACTCAGACAGCGGATTTCAGATTTCGAGTGTATTCG 2397

DB 4606 GATTTCAGACTCCGACAGCGGATTTCAGATTTCAGACAGCGGACTCAGATTTCGATAGTGTATTC 4665
QY 2398 GACTCAGATAGCGATTCCGACTCAGATAGCGGACTCAGATTTCAGACAGCGGATTTCAGATTCA 2457
DB 4666 GATTTCAGACAGTACTCGGATTTCGATAGTGTACTCAGATTTCAGACAGTGTACTCAGATTCA 4725
QY 2458 GACAGCGATTTCAGATTTCAGATAGCGGATTTCGATAGTGTACTCAGATTTCGATAGTGTACT 2517
DB 4726 GATAGCGACTCAGATTTCAGACAGTGTATTCGAGTGTACTCAGATTTCGATAGTGTACTCAG 4785
QY 2518 GACTCGGATTTCAGATAGCGGATTTCAGATTTCGATAGTGTACTCAGATTTCGATAGTGTACT 2577
DB 4786 GATTTCGATTTCGATAGCGGATTTCGATAGTGTACTCAGATTTCGATAGTGTACTCAGATTTCG 4845
QY 2578 GACTCAGACAGTGTATTCGAGTGTACTCAGCGAGTGTATTCGATTTCAGATAGTGTATTCGACTCC 2637
DB 4846 GACTCAGACAGCGACTCCGATTTCAGATAGTGTATTCGATAGTGTACTCAGACAGCGGATTTCG 4905
QY 2638 GACAGTGTACTCGGATTTCAGATAGCGGACTCAGACTCCGATAGCGGACTCCGATTTCAGATAGC 2697
DB 4906 GATAGTGTACTCGGATTTCAGACAGTGTATTCGAGTGTACTCAGACAGCGGACTCCGATTTCAG 4965
QY 2698 GATTTCGACTCAGATAGCGGATTTCAGATAGCGGATTTCAGATAGCGGATTTCAGACAGCGGATTTC 2757
DB 4966 GATTTCGAGCGGACAGACAGCGGACTCCGATTTCAGATAGTGTATTCGAGCGGACGACAGCGGATTTC 5025
QY 2758 GATTTCAGACAGCGGACTCAGACAGTGTACTCAGATTTCAGATAGTGTACTCAGGATTTCAGCGGAGT 2817
DB 5026 GATAGTGTACTCGGATTTCAGACAGTGTATTCGAGTGTACTCAGACAGCGGATTTCGAGTGTACT 5085
QY 2818 GATTTCAGACTCAGGTAGTGTACTCCGATTTCATCAAGTGTATTCGAGTGTACTCAGAAAGTGTATTC 2877
DB 5086 GACTTCGACTCAGATAGCGGATTTCGATTTCAGACAGTGTACTCAGGACTCAGATAGTGTACTCC 5145
QY 2878 AATGAGCGATTTCGAGTGTACTCAGGTTCTTAAACAATATGTA 2913
DB 5146 GATTTCAGACAGCGGACTCCGATTTCGATAGTGTATTCGATAGTGTATTCGATAGTGTATTCG 5181

RESULT 15
AAH54581
ID AAH54581 standard; DNA; 4627 BP.
XX
AC AAH54581;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3945.
XX Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
KW endocarditis; ds.
XX Staphylococcus epidermidis.
QS WO200134809-A2.
PN 17-MAY-2001.
XX
PD 09-NOV-2000; 2000WO-US030782.
HF 09-NOV-1999; 99US-0164258P.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Zimmerly WJ;
XX WPI; 2001-316495/33.
XX
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
XX
XX Claim 8; Page 1606-1608; 2188pp; English.
XX

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC *epidermidis* polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to
CC AAH55098 represent specifically claimed *S. epidermidis* genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX
SQ Sequence 4627 BP; 1480 A; 849 C; 1077 G; 1221 T; 0 U; 0 Other;

Query Match 16.9%; Score 589.6; DB 4; Length 4627;

Best Local Similarity 71.9%; Pred. No. 7.4e-97;

Matches 817; Conservative 0; Mismatches 304; Indels 15; Gaps 3;

QY 1978 GATTTCAGATTCTGACCCAGGTTTCAGATTCTGCGAGCGATTCTTAATTCAGATAGCGGTTCA 2037

DB 3454 GACTCCGATTTCAGATAGTGGATTCGGATGACAGAGCGACTCGGATTCGACAGTGAATCG 3513

QY 2038 GATTCGGGTAGTGATTCATCATCAGATAGTGGTTCAGATTTCAGCGAGTGATTCAGATTCA 2097

DB 3514 GACGACAGACAGTGACTCAGATTTCAGACAGTGTTCGGACGCGACAGCGACTCGGATTC 3573

QY 2098 GCAAGTGATTCAGATCGAGGAGTGATTCAGATTTCAGCAGACGCGATTCGAGCTCAGCGAGC 2157

DB 3574 GATAGTGATTCGGACGCGACAGAGTGACTCAGATTTCAGACAGTGACTCGGATTCAGACAGC 3633

QY 2158 -----GATTCGGACTCAGACAACTGACTCGGATTCAGATAGCGATTCTGACTCAGACAGT 2211

DB 3634 GATTCGGATTTCGATTTCAGACAGTGACTCGGATTCAGACAGTGACTCAGACTCCGACAGT 3693

QY 2212 GACTCAGATTTCGACAGTGACTCAGATTTCAGATAGCGATTCTGACTCAGACAGTGACTCA 2271

DB 3694 GATTCGGATTTCAGATAGCGACTCAGACGCGAGATAGTGATTCGATTTCAGATAGCGACTCC 3753

QY 2272 GATTCAGATAGCGGATTTCAGATTTCAGATAGCGGATTCAGATTTCGACAGTGATTCCGACTCA 2331

DB 3754 GACGACAGATAGTGATTTCGGACGCGACAGAGTGACTCAGATTTCATACAGTGACTCAGATTCA 3813

QY 2332 GACAGCGATTTCGACTCCGACAGTGATTTCGGACTCAGACAGCGATTTCAGATTTCGACAGT 2391

DB 3814 GACAGTGATTTCGGACGCGACAGTGACTCCGACTCCGACGCGATTTCAGACTCAGATAGT 3873

QY 2392 GATTCGGACTCAGATAGCGATTTCGAGTTTCAGATAGCGACTCAGATTTCAGACAGCGATTCA 2451

DB 3874 GACTCAGACGCGACAGTGACTCGGACTCAGATTTCAGATAGTGATTTCAGATGCGAAGCGATTCA 3933

QY 2452 GATTCAGACAGCGGATTTCAGATTTCAGATAGCGGATTTCAGATTTCGACAGTGACTCAGATTCC 2511

DB 3934 GACTCAGATAGCGACTCCGATTTCAGACAGCGACTCCGACTCAGACAGTGATTTCGATTTC 3993

QY 2512 GACAGTGACTCGGATTTCAGATAGCGGATTTCAGATTTCGACAGTGACTCAGATTTCGACAGT 2571

DB 3994 GACAGCGATTTCGACTCAGATAGTGACTCAGACGCGAGATAGTGATTTCGATTTCAGATAGC 4053

QY 2572 GACTCAGACTCAGACAGTGATTTCGGATTTCAGGAGTGATTTCGGATTTCAGATAGTGATTCC 2631

DB 4054 GACTCCGATTTCGATAGTGACTCCGATTTCAGATAGCGACTCCGATTTCAGATAGTGATTTCG 4113

QY 2632 GACTCCGACAGTGACTCGGATTTCAGATAGCGACTCAGACTCGGATAGCGACTCGGATTCA 2691

|||

DB 4114 GACGACAGACAGTGACTCGTACTCAGATAGTGACTCGGATTTCAGACAGTGATTTCGGATTCC 4173

QY 2692 GATACCGGATTTCGGACTCAGATAGCGGATTTCAGAAATCAGACAGCGATTTCAGAAATCAGACAGC 2751

DB 4174 GATACCGGATTTCGGATTTCGGATAGTGACTCGGATTTCAGACAGTGATTTCGGACTCAGACAGC 4233

QY 2752 GATTTCAGATTTCAGACAG-----CGACTCAGACAGTGACTCAGATTTCAGATAGTGACTCG 2805

DB 4234 GACTCCGATTTCAGATAGTGATTTCGGACTCAGACAGCGATTTCGGATTTCGATAGTGACTCG 4293

QY 2806 GATTTCAGCGAGTGATTTCAGACTCAGGTTAGTGACTCCGATTTCATCAAGTGATTTCGACTCA 2865

DB 4294 GATTTCAGACAGTGATTTCGGACTCAGACAGCGGATTTCGGATTTCGATAGTGACTCAGATTCA 4353

QY 2866 GAAAGTGATTTCAAATAGCGGATTTCGGAGTCAGGTTCTTAACAATAATAGTAGTTCCGCTTAAT 2925

DB 4354 GACAGTGATTTCGGGCTCAGACAGCGGATTTCGGATTTCAGACAGTGACTTCGGACTCAGATTAGT 4413

QY 2926 TCACCTAAAAATGGTGACTTAATGCTTCTTAATAAAAAATGAGGCTAAAGATAGTAAAGAACCA 2985

DB 4414 GACTCCGATTTCAGACAGCGGACTCGGATTTCGTATAAAAAATGCAAAAGA---TAAATTACCT 4470

QY 2986 TTACCAGATACAGGTTCTGAAGATGAACAAATAGCTCACTAAATTTGGGGATTATTAGCA 3045

DB 4471 GATACAGGAGCAAAATGAAGATCATGATTCTTAAGGCGACATTACTTTGGAACCTTTATTTCGA 4530

QY 3046 TCAATAGGTTTCATTACTACTTTTCAGAGAAGAAAAAGAAAAATAAGATAAGAAAAATA 3101

DB 4531 GGTTTAGGAGCATTTATTATTAGGAGAGCTCGTAAAAAAGATATAAAGAAAAATA 4586

Search completed: November 5, 2004, 06:04:58

Job time : 1579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 05:38:03 ; Search time 276 Seconds
(without alignments)
9008.476 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaccataataacacatc.....gaagagtataaagaagctt 3498

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3498	100.0	3498	3	US-08-293-728-1
2	3498	100.0	3498	3	US-09-421-868-1
3	3399.6	97.2	4709	4	US-08-956-171E-587
4	3399.6	97.2	4709	4	US-08-956-171E-587
5	619.2	17.7	5406	4	US-08-781-986A-587
6	619.2	17.7	5406	4	US-08-386-962C-1
7	589.6	16.9	4627	4	US-09-386-959-1
8	584.8	16.7	1485	4	US-09-710-279-35
9	568.6	16.3	2969	4	US-09-200-650E-2
10	510	14.6	11050	4	US-08-956-171E-96
11	510	14.6	11050	4	US-08-781-986A-96
12	462	13.2	3600	4	US-09-147-405B-14
13	448.4	12.8	4358	4	US-08-956-171E-454
14	448.4	12.8	4358	4	US-08-781-986A-454
15	436.8	12.5	756	4	US-09-710-279-31
16	365	10.4	396	4	US-08-956-171E-3602
17	365	10.4	396	4	US-08-781-986A-3602
18	360.6	10.3	3498	4	US-09-200-650E-8
19	360.6	10.3	4005	4	US-08-956-171E-631
20	360.6	10.3	4005	4	US-08-781-986A-631
21	359.8	10.3	2841	4	US-09-200-650E-4
22	340.8	9.7	864	4	US-09-710-279-467
23	340.8	9.7	3246	4	US-09-710-279-3336
24	340.8	9.7	3619	4	US-09-710-279-4157
25	327.6	9.4	399	4	US-08-956-171E-3555
26	327.6	9.4	399	4	US-08-781-986A-3555
27	307.6	8.8	1461	4	US-09-710-279-787

c	28	307.6	8.8	3612	4	US-09-710-279-3696	Sequence 3696, Ap
	29	303.4	8.7	400	4	US-08-956-171E-4206	Sequence 4206, Ap
	30	303.4	8.7	400	4	US-08-781-986A-4206	Sequence 4206, Ap
c	31	302.2	8.6	530	4	US-08-956-171E-3656	Sequence 3656, Ap
c	32	302.2	8.6	530	4	US-08-781-986A-3656	Sequence 3656, Ap
	33	300.8	8.6	1464	4	US-09-386-962C-13	Sequence 13, Appl
	34	300.8	8.6	1464	4	US-09-386-959-64	Sequence 64, Appl
	35	292.4	8.4	3945	4	US-09-200-650E-6	Sequence 6, Appli
c	36	292.2	8.4	573	4	US-08-956-171E-3859	Sequence 3859, Ap
c	37	292.2	8.4	573	4	US-08-781-986A-3859	Sequence 3859, Ap
c	38	272.4	7.8	476	4	US-08-956-171E-3725	Sequence 3725, Ap
c	39	272.4	7.8	476	4	US-08-781-986A-3725	Sequence 3725, Ap
	40	269.2	7.7	374	4	US-08-956-171E-4060	Sequence 4060, Ap
	41	269.2	7.7	374	4	US-08-781-986A-4060	Sequence 4060, Ap
	42	260.8	7.5	308	4	US-08-956-171E-4466	Sequence 4466, Ap
	43	260.8	7.5	308	4	US-08-781-986A-4466	Sequence 4466, Ap
c	44	250.8	7.2	400	4	US-08-956-171E-3600	Sequence 3600, Ap
c	45	250.8	7.2	400	4	US-08-781-986A-3600	Sequence 3600, Ap

ALIGNMENTS

RESULT 1
US-08-293-728-1
; Sequence 1, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3498
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (301)..(3099)
US-08-293-728-1

Query Match	100.0%;	Score 3498;	DB 3;	Length 3498;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3498;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGTACCATAAATACACATCTGCTTTTGA	AAATATGATTTTCAAGCTAGGATACATTA	60
Db	1	GGTACCATAAATACACATCTGCTTTTGA	AAATATGATTTTCAAGCTAGGATACATTA	60
QY	61	GGTAGAGTTTCATATTAATAAATAAATGTTT	GCAATCGATCGTGTGCTTTGTA	120
Db	61	GGTAGAGTTTCATATTAATAAATAAATGTTT	GCAATCGATCGTGTGCTTTGTA	120
QY	121	ATTCCTTAATAGCAATAAATAAATGTTTGT	TAGTAAGTATTTTGGATTAATAAAA	180
Db	121	ATTCCTTAATAGCAATAAATAAATGTTTGT	TAGTAAGTATTTTGGATTAATAAAA	180
QY	181	TATCGATACAAATTAATTCGATTAATTCG	ATTTAGTATTAATTCATTAACAGAGATT	240
Db	181	TATCGATACAAATTAATTCGATTAATTCG	ATTTAGTATTAATTCATTAACAGAGATT	240
QY	241	AAATATATCTTAAGGGTATATAGTTAAATA	AAATGACCTTTTAAAGAGGGAATAAAA	300
Db	241	AAATATATCTTAAGGGTATATAGTTAAATA	AAATGACCTTTTAAAGAGGGAATAAAA	300
QY	301	ATGAATATGAAGAAAAAAGAAAAACGCAAT	TTCGAAAAAATCGATTCGCGTGGCTTCA	360
Db	301	ATGAATATGAAGAAAAAAGAAAAACGCAAT	TTCGAAAAAATCGATTCGCGTGGCTTCA	360

QY 361 GTGCTGTAGGTACGTTAAATCGTTTTCGACTACTCAGCAGTAAAGAAAGCAGATCAAGT 420
Db |||||
QY 361 GTGCTGTAGGTACGTTAAATCGTTTTCGACTACTCAGCAGTAAAGAAAGCAGATCAAGT 420
Db |||||
QY 421 GAAATAGTGTTCGCAATCTGATAGCGCAAGTAAAGAAAGCAGATCAAGT 480
Db |||||
QY 421 GAAATAGTGTTCGCAATCTGATAGCGCAAGTAAAGAAAGCAGATCAAGT 480
Db |||||
QY 481 AGCGTGTAGTGTGCACTTAAACAGACGACACAAACGTCGATCTAAACATCGTCA 540
Db |||||
QY 481 AGCGTGTAGTGTGCACTTAAACAGACGACACAAACGTCGATCTAAACATCGTCA 540
Db |||||
QY 541 AACACTAATATGCGGAAACGAGTGTGCGCAAAATCCAGCAACAGGAAACGACACA 600
Db |||||
QY 541 AACACTAATATGCGGAAACGAGTGTGCGCAAAATCCAGCAACAGGAAACGACACA 600
Db |||||
QY 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGCTGTAAGTACTACTAG 660
Db |||||
QY 601 TCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACCTGCTGTAAGTACTACTAG 660
Db |||||
QY 661 ACAAGCAATCAAGTAAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGAA 720
Db |||||
QY 661 ACAAGCAATCAAGTAAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGAA 720
Db |||||
QY 721 TTAGTGAATCAAAACAGTAAATGAAACGACTTTTAAATGATCACTAAATCACTGTA 780
Db |||||
QY 721 TTAGTGAATCAAAACAGTAAATGAAACGACTTTTAAATGATCACTAAATCACTGTA 780
Db |||||
QY 781 AATTCACCTCAAAATTTCAAAATCGGAAATGTTTCAACACGCAAGATCTTCAACT 840
Db |||||
QY 781 AATTCACCTCAAAATTTCAAAATCGGAAATGTTTCAACACGCAAGATCTTCAACT 840
Db |||||
QY 841 GAAGCAACACCTTCAAAATGAAATCAAGTCCAGAGTACAGATGCAAGTAAAGAT 900
Db |||||
QY 841 GAAGCAACACCTTCAAAATGAAATCAAGTCCAGAGTACAGATGCAAGTAAAGAT 900
Db |||||
QY 901 GTAGTAAATCAAGCGGTTAAATCAAGTGGCGCTAGATGAGAGCAATTTAGTTAGCGCA 960
Db |||||
QY 901 GTAGTAAATCAAGCGGTTAAATCAAGTGGCGCTAGATGAGAGCAATTTAGTTAGCGCA 960
Db |||||
QY 961 GTAGTGCAGATGCAACCGGAGTGGCAAGATATACGAATCAGTTGACGAATGTGACA 1020
Db |||||
QY 961 GTAGTGCAGATGCAACCGGAGTGGCAAGATATACGAATCAGTTGACGAATGTGACA 1020
Db |||||
QY 1021 GTTGTATTTGACTCTGTGACGATGATCGCATCGCAAGAGTATGTCGCAACTGAT 1080
Db |||||
QY 1021 GTTGTATTTGACTCTGTGACGATGATCGCATCGCAAGAGTATGTCGCAACTGAT 1080
Db |||||
QY 1081 TATGTTTTCAGTGCCTAATTTCTGCTTTAAAGGTGACACATTCAAATTAACCTGTACCT 1140
Db |||||
QY 1081 TATGTTTTCAGTGCCTAATTTCTGCTTTAAAGGTGACACATTCAAATTAACCTGTACCT 1140
Db |||||
QY 1141 AAAGAAATTAACCTTAAATCGGTGTAATCTCACTGCTTAAAGTGCACCAATTTATGCGTGA 1200
Db |||||
QY 1141 AAAGAAATTAACCTTAAATCGGTGTAATCTCACTGCTTAAAGTGCACCAATTTATGCGTGA 1200
Db |||||
QY 1201 GATCAAGTATTCGCAAAATCGGTGTAATCGATGATGATGATGATGATGATGATGATGATGAT 1260
Db |||||
QY 1201 GATCAAGTATTCGCAAAATCGGTGTAATCGATGATGATGATGATGATGATGATGATGATGAT 1260
Db |||||
QY 1261 GACTATGTAATTAACCTTAAAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1320
Db |||||
QY 1261 GACTATGTAATTAACCTTAAAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1320
Db |||||
QY 1321 CTTGAAATGTTTAAAGACAGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1380
Db |||||
QY 1321 CTTGAAATGTTTAAAGACAGTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1380
Db |||||
QY 1381 GCAACAAACAGATTTAGTATGATTAAGAAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1440
Db |||||
QY 1381 GCAACAAACAGATTTAGTATGATTAAGAAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1440
Db |||||
QY 1441 AAAGGTACAATTTGACCAAAATCGATAAAACAAATTAACGTATCGTCAGACAAATTTATGTC 1500

Db 1441 AAAGGTACAATTTGACCAAAATCGATAAAACAAATTAACGTATCGTCAGACAAATTTATGTC 1500
QY 1501 AATCCAAGTGCAGATTAACGTTTATTCGCGCGGTTTTAAACAGTAAATTTAAACCAATACG 1560
Db |||||
QY 1501 AATCCAAGTGCAGATTAACGTTTATTCGCGCGGTTTTAAACAGTAAATTTAAACCAATACG 1560
Db |||||
QY 1561 GATAGTAATGATTAATTAATAGATCAGCAAAATACAAATTAATAAGTATATAAAGTAGATAAT 1620
Db |||||
QY 1561 GATAGTAATGATTAATTAATAGATCAGCAAAATACAAATTAATAAGTATATAAAGTAGATAAT 1620
Db |||||
QY 1621 GCAGCTGATTTATCTGAAAAGTTTACTTTGTAATCCAGAAAACCTTTGAGGATGCTCAAT 1680
Db |||||
QY 1621 GCAGCTGATTTATCTGAAAAGTTTACTTTGTAATCCAGAAAACCTTTGAGGATGCTCAAT 1680
Db |||||
QY 1681 AGTGTGAATATTAACATTCGCAAAATCCAAATCAATTAAGTAGAGTAAATACGCGCTGAT 1740
Db |||||
QY 1681 AGTGTGAATATTAACATTCGCAAAATCCAAATCAATTAAGTAGAGTAAATACGCGCTGAT 1740
Db |||||
QY 1741 GATCAAAATTAACAAACCGGTATATAGTATGTTTAAATGCTATATGATCCGAAATAGCAAA 1800
Db |||||
QY 1741 GATCAAAATTAACAAACCGGTATATAGTATGTTTAAATGCTATATGATCCGAAATAGCAAA 1800
Db |||||
QY 1801 GGTGATTTAGCTTTACGTTTCACTTTATATGCGGTAAATCTGAAATATATTTGGCGCTCT 1860
Db |||||
QY 1801 GGTGATTTAGCTTTACGTTTCACTTTATATGCGGTAAATCTGAAATATATTTGGCGCTCT 1860
Db |||||
QY 1861 ATGTCATCGGCAACGAGTAGTCAATTAACCGATCAGGTTCTGCTGACGCTATCGAT 1920
Db |||||
QY 1861 ATGTCATCGGCAACGAGTAGTCAATTAACCGATCAGGTTCTGCTGACGCTATCGAT 1920
Db |||||
QY 1921 AAACAGTTGTTTCCTGAAACAACTGATGAGCTGCTGAAATGAAACCAATTCAGAGGAT 1980
Db |||||
QY 1921 AAACAGTTGTTTCCTGAAACAACTGATGAGCTGCTGAAATGAAACCAATTCAGAGGAT 1980
Db |||||
QY 1981 TCAGATTTGACCCAGGTTTCAGATTTGCGAGCGATTTCAATTCAGATAGCGGTTTCAGAT 2040
Db |||||
QY 1981 TCAGATTTGACCCAGGTTTCAGATTTGCGAGCGATTTCAATTCAGATAGCGGTTTCAGAT 2040
Db |||||
QY 2041 TCGGCTAGTGTATTCACATCAGATAGTGTGTTTCAAGTTCAGAGTGTGATTCAGATTCAGCA 2100
Db |||||
QY 2041 TCGGCTAGTGTATTCACATCAGATAGTGTGTTTCAAGTTCAGAGTGTGATTCAGATTCAGCA 2100
Db |||||
QY 2101 AGTGATTCAGACTCAGCGAGTGTGATTCAGATTCAGCAAGCGATTCGACTCAGCGAGCGAT 2160
Db |||||
QY 2101 AGTGATTCAGACTCAGCGAGTGTGATTCAGATTCAGCAAGCGATTCGACTCAGCGAGCGAT 2160
Db |||||
QY 2161 TCCGACTCAGACAATGATTCGGATTCAGATAGCGATTCGACTCAGACAAGTCACTCAGAT 2220
Db |||||
QY 2161 TCCGACTCAGACAATGATTCGGATTCAGATAGCGATTCGACTCAGACAAGTCACTCAGAT 2220
Db |||||
QY 2221 TCCGACAGTGTACTCAGATTCAGATAGCGATTCGACTCAGACAAGTCACTCAGATTCAGAT 2280
Db |||||
QY 2221 TCCGACAGTGTACTCAGATTCAGATAGCGATTCGACTCAGACAAGTCACTCAGATTCAGAT 2280
Db |||||
QY 2281 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGTTCGACTCAGACAAGCGAT 2340
Db |||||
QY 2281 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGTTCGACTCAGACAAGCGAT 2340
Db |||||
QY 2341 TCTGACTCGACAGTGTATTCGACTCAGACAAGCGATTCAGATTCGCAAGTGTTCGACTCAGACAAG 2400
Db |||||
QY 2341 TCTGACTCGACAGTGTATTCGACTCAGACAAGCGATTCAGATTCGCAAGTGTTCGACTCAGACAAG 2400
Db |||||
QY 2401 TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAAGCGATTCAGATTCAGAT 2460
Db |||||
QY 2401 TCAGATAGCGATTCGACTCAGATAGCGATTCAGATTCAGACAAGCGATTCAGATTCAGAT 2460
Db |||||
QY 2461 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGTTCGACTCAGACAAGCGAT 2520
Db |||||
QY 2461 AGCGATTCAGATTCAGATAGCGATTCAGATTCGCAAGTGTTCGACTCAGACAAGCGAT 2520
Db |||||
QY 2521 TCGGATTCAGATAGCGATTCAGATTCGCAAGTGTTCGACTCAGATTCGCAAGTGTTCGACTCAGAC 2580

QY	781	AATTCACTCAAAATTTCTACAAATCGGAATGTTTCAACAACGCAAGATACCTCAACT	840	QY	1861	ATGTCATGGGCAACAAAGTAGTAGCTTAATTAACCGATCAGGTTCTGGTACGGTATCGAT	1920
Db	781	AATTCACTCAAAATTTCTACAAATCGGAATGTTTCAACAACGCAAGATACCTCAACT	840	Db	1861	ATGTCATGGGCAACAAAGTAGTAGCTTAATTAACCGATCAGGTTCTGGTACGGTATCGAT	1920
QY	841	GAAGCAACACCTTTCAAAACAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT	900	QY	1921	AAACAGATTGTTCTCTGAAACAACTGATGAGCTGGTGAAATTTGAACCAATTTCCAGAGAT	1980
Db	841	GAAGCAACACCTTTCAAAACAATGAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGAT	900	Db	1921	AAACAGATTGTTCTCTGAAACAACTGATGAGCTGGTGAAATTTGAACCAATTTCCAGAGAT	1980
QY	901	GTAGTTAATCAAGCGGTTAATCAAGTGGGCTAGAAATGAGAGCAATTTAGTTAGCGCA	960	QY	1981	TCAGATTTCTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCAGAT	2040
Db	901	GTAGTTAATCAAGCGGTTAATCAAGTGGGCTAGAAATGAGAGCAATTTAGTTAGCGCA	960	Db	1981	TCAGATTTCTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCAGAT	2040
QY	961	GTAGCTGAGATGCAACCGGAGCTGGCAGAGATATACGAATCAGTTGACCGAATGTGACA	1020	QY	2041	TCGGGTAGTGTATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA	2100
Db	961	GTAGCTGAGATGCAACCGGAGCTGGCAGAGATATACGAATCAGTTGACCGAATGTGACA	1020	Db	2041	TCGGGTAGTGTATTTCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCA	2100
QY	1021	GTGGTATTTGACTCTGGTACGACTGTGATCCGCCCAAGCAGGTTATGTCAAACTGAAT	1080	QY	2101	AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCAAGCGATTTCCGACTCAGCGAGCGAT	2160
Db	1021	GTGGTATTTGACTCTGGTACGACTGTGATCCGCCCAAGCAGGTTATGTCAAACTGAAT	1080	Db	2101	AGTGATTCAGACTCAGCGAGTGAATTCAGATTCAGCAAGCGATTTCCGACTCAGCGAGCGAT	2160
QY	1081	TATGGTTTTTTCAGTGCCTAAATTCCTGCTGTTAAAGTGACACATTCAAAATTAATCTGACT	1140	QY	2161	TCCGACTCAGACAATGACTTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGAT	2220
Db	1081	TATGGTTTTTTCAGTGCCTAAATTCCTGCTGTTAAAGTGACACATTCAAAATTAATCTGACT	1140	Db	2161	TCCGACTCAGACAATGACTTCGGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGAT	2220
QY	1141	AAAGAAATTAACCTTAAATGGTGTAATTCCTCACTGCTAAAGTGCCACCAATTAATGGCTGA	1200	QY	2221	TCCGACAGTGAATTCAGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT	2280
Db	1141	AAAGAAATTAACCTTAAATGGTGTAATTCCTCACTGCTAAAGTGCCACCAATTAATGGCTGA	1200	Db	2221	TCCGACAGTGAATTCAGATTCAGATAGCGATTTCTGACTCAGACAGTGAATTCAGATTCAGAT	2280
QY	1201	GATCAAGTATTTGGCAAAATGGTGTAATTCGATAGTGAATGTTAATTTATATACATTTACA	1260	QY	2281	AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGAATTCGGACTCAGACAGCGAT	2340
Db	1201	GATCAAGTATTTGGCAAAATGGTGTAATTCGATAGTGAATGTTAATTTATATACATTTACA	1260	Db	2281	AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGAATTCGGACTCAGACAGCGAT	2340
QY	1261	GACTATGTAATTAAGATGATGTAAGAAAGCACTTTGACCATCGCCGCTTATATGAC	1320	QY	2341	TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGACAGTGAATTCGAC	2400
Db	1261	GACTATGTAATTAAGATGATGTAAGAAAGCACTTTGACCATCGCCGCTTATATGAC	1320	Db	2341	TCTGACTCCGACAGTGAATTCGACTCAGACAGCGATTCAGATTCGACAGTGAATTCGAC	2400
QY	1321	CCTGAAATGTTTAAAGACAGTGAATGTGACATTTGGCTACTCGCATAGGTAGTACAACA	1380	QY	2401	TCAGATAGCGATTCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGAT	2460
Db	1321	CCTGAAATGTTTAAAGACAGTGAATGTGACATTTGGCTACTCGCATAGGTAGTACAACA	1380	Db	2401	TCAGATAGCGATTCGACTCAGATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGAT	2460
QY	1381	GCAACAAACAGTATAGTAGATTTGAAAATATGGTAAAGTTTATAAATTTATCTATT	1440	QY	2461	AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGAATTCGGACTCAGATTCGACAGTGA	2520
Db	1381	GCAACAAACAGTATAGTAGATTTGAAAATATGGTAAAGTTTATAAATTTATCTATT	1440	Db	2461	AGCGATTCAGATTCAGATAGCGATTTCCGACAGTGAATTCGGACTCAGATTCGACAGTGA	2520
QY	1441	AAAGTCAAAATGACCAATCGATTAACCAATTAATGATGCTGACAGCAATTTATGTC	1500	QY	2521	TCGATTCAGATAGCGATTCGACTCAGATTCGACAGTGAATTCGACAGTGAATTCGAC	2580
Db	1441	AAAGTCAAAATGACCAATCGATTAACCAATTAATGATGCTGACAGCAATTTATGTC	1500	Db	2521	TCGATTCAGATAGCGATTCGACTCAGATTCGACAGTGAATTCGACAGTGAATTCGAC	2580
QY	1501	AATCCAAAGTGAATGAACTTATTCGCGCGGTTTAAACAGGTAATTTAAACCAAAATACG	1560	QY	2581	TCAGACAGTGAATTCGATTCAGCGAGTGAATTCGATTCAGATTCGATTCGATTCGATTCG	2640
Db	1501	AATCCAAAGTGAATGAACTTATTCGCGCGGTTTAAACAGGTAATTTAAACCAAAATACG	1560	Db	2581	TCAGACAGTGAATTCGATTCAGCGAGTGAATTCGATTCAGATTCGATTCGATTCGATTCG	2640
QY	1561	GATAGTAATGCAATTAATAGTACAGCAAAATACAGTATTAAGTATATAAAGTAGATAAT	1620	QY	2641	AGTGACTTCGGATTCAGATAGCGACTCAGACTCGGATTCGGATTCGGATTCAGATAGCGAT	2700
Db	1561	GATAGTAATGCAATTAATAGTACAGCAAAATACAGTATTAAGTATATAAAGTAGATAAT	1620	Db	2641	AGTGACTTCGGATTCAGATAGCGACTCAGACTCGGATTCGGATTCGGATTCAGATAGCGAT	2700
QY	1621	GCAGCTGATTTATCTGAAAGTACTTTGTGATCCGAAACCTTTGAGGATGCTCACTAAT	1680	QY	2701	TCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGAT	2760
Db	1621	GCAGCTGATTTATCTGAAAGTACTTTGTGATCCGAAACCTTTGAGGATGCTCACTAAT	1680	Db	2701	TCGACTCAGATAGCGATTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGAT	2760
QY	1681	AGTGATGATATTTACATTTCCCAATCCAAATCAATTAAGTATAGTTTAAATACGCTGAT	1740	QY	2761	TCAGACAGCGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG	2820
Db	1681	AGTGATGATATTTACATTTCCCAATCCAAATCAATTAAGTATAGTTTAAATACGCTGAT	1740	Db	2761	TCAGACAGCGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG	2820
QY	1741	GATCAAAATTAACACCGGATATAGTAGTTGTTAATGCTATTTGATCCGAAATAGCAAA	1800	QY	2821	TCAGACTCAGGTAGTGAATTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG	2880
Db	1741	GATCAAAATTAACACCGGATATAGTAGTTGTTAATGCTATTTGATCCGAAATAGCAAA	1800	Db	2821	TCAGACTCAGGTAGTGAATTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG	2880
QY	1801	GGTGATTTAGCTTTTACGTTCACTTTATATGGGTATTAATCGAATATATTTGGCGCTCT	1860	QY	2881	AGCGATTCGAGTCAAGGTTCTAAACAATTAATGTAGTTCGCGCTTAATTCACCTAAAAATG	2940
Db	1801	GGTGATTTAGCTTTTACGTTCACTTTATATGGGTATTAATCGAATATATTTGGCGCTCT	1860	Db	2881	AGCGATTCGAGTCAAGGTTCTAAACAATTAATGTAGTTCGCGCTTAATTCACCTAAAAATG	2940
				QY	2941	ACTAATGCTTCTAATAAAAAATGAGGCTAAAGATAGTAAGAACCATTTACCATACAGAT	3000

2941 ACTAATGCTTCTTAATAAATAGGCTAAAGATAGTAAGAACCAATACAGATACAGGT 3000
3001 TCTGAGATGAAGCAATACGCTACTAATTTGGGATATTAGCATCAATAGGTTCAITA 3060
3001 TCTGAGATGAAGCAATACGCTACTAATTTGGGATATTAGCATCAATAGGTTCAITA 3060
3061 CTACTTTTCAGAGCAAAAAAGAAAAAATAAGATAAGAAATAGTAATTAATTAAT 3120
3061 CTACTTTTCAGAGCAAAAAAGAAAAAATAAGATAAGAAATAGTAATTAATTAAT 3120
3121 TAATCATATGATTCATGAAGAGCCACCTTAAAGAGGTCCTTTTACTTTGGATTTTCCA 3180
3121 TAATCATATGATTCATGAAGAGCCACCTTAAAGAGGTCCTTTTACTTTGGATTTTCCA 3180
3181 AATATATGTTTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240
3181 AATATATGTTTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240
3241 TAGATGTTTATATATTTGGCTTGGCGAAAAAATAAGGTTGAAGTAGGTTGTTAATTAGG 3300
3241 TAGATGTTTATATTTGGCTTGGCGAAAAAATAAGGTTGAAGTAGGTTGTTAATTAGG 3300
3301 GAAAAATTAAGAGCAAAAAATACAGTTGAAAAATTAATTTCTAGTTTGAAGTAGG 3360
3301 GAAAAATTAAGAGCAAAAAATACAGTTGAAAAATTAATTTCTAGTTTGAAGTAGG 3360
3361 ATGTTATCAAAATTTGGGAAAGTATCGTGGAGTGCAGTGGTTCTCGGGAGAGAA 3420
3361 ATGTTATCAAAATTTGGGAAAGTATCGTGGAGTGCAGTGGTTCTCGGGAGAGAA 3420
3421 TCCATATGATCTGAGTCGTTGAAACTGACTAATAATAAAAAATAAACTTAGAACAGTAGA 3480
3421 TCCATATGATCTGAGTCGTTGAAACTGACTAATAATAAAAAATAAACTTAGAACAGTAGA 3480
3481 AGAGTATAAGAAAGCTT 3498
3481 AGAGTATAAGAAAGCTT 3498

RESULT 3
US-08-956-171E-587
; Sequence 587, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956.171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PE248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 587:
SEQUENCE CHARACTERISTICS:
LENGTH: 4709 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 587:
US-08-956-171E-587
Query Match 97.2%; Score 3399.6; DB 4; Length 4709;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3466; Conservative 1; Mismatches 11; Indels 21; Gaps 4;
QY 1 GGTACCAATAAATACACATCTGCTTTTGAATAAATATGATTTCAGCTAGGATTACATTA 60
Db 1204 GGTACCAATAAATACACATCTGCTTTTGAATAAATATGATTTCAGCTAGGATTACATTA 1263
QY 61 GGTAGAGTTCATATTAATAAATAAATGTTTGCATCAATCGTACGTTGTCGTTGTA 120
Db 1264 GGTAGAGTTCATATTAATAAATAAATGTTTGCATCAATCGTACGTTGTCGTTGTA 1323
QY 121 ATTCCTTAAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA 180
Db 1324 ATTCCTTAAAAATAGCAATAAATAAATGTTTGTAGTAAAGTATTATTGTGGATAATAAAA 1383
QY 181 TATCGATACAAATTAATGCTATTAATGCAATTTTGTAGTATTAATTCATTAACAGAGATT 240
Db 1384 TATCGATACAAATTAATGCTATTAATGCAATTTTGTAGTATTAATTCATTAACAGAGATT 1443
QY 241 AATATATATC-TTAAAGGTTATATAGTTAATATAAATAAGTCTTTTAAATAAGAGGATAA 299
Db 1444 AATATATATC-TTAAAGGTTATATAGTTAATATAAATAAGTCTTTTAAATAAGAGGATAA 1503
QY 300 AATGAATATGAAGAAAAAAGAAAAACACCAATTCGGAATAAATCGATTGCGGTGCTTC 359
Db 1504 AATGAATATGAAGAAAAAAGAAAAACACCAATTCGGAATAAATCGATTGCGGTGCTTC 1563
QY 360 AGTGCTTTGAGGTACGTTAATCGGTTTGGACATCTCAGCAGTAAGAAAGACAGATGCAAG 419
Db 1564 AGTGCTTTGAGGTACGTTAATCGGTTTGGACATCTCAGCAGTAAGAAAGACAGATGCAAG 1623
QY 420 TGAATAATAGTTTACGCAATCTGATAGCGCAAGTAAGCAAGCAAGCAAGTAATGATTCAAG 479
Db 1624 TGAATAATAGTTTACGCAATCTGATAGCGCAAGTAAGCAAGCAAGTAATGATTCAAG 1683
QY 480 TAGCGTTTAGTCTGCACCTTAAACACAGACACAAACGTCGAGTGATCTACTAAAAACATGTC 539
Db 1684 TAGCGTTTAGTCTGCACCTTAAACACAGACACAAACGTCGAGTGATCTACTAAAAACATGTC 1743
QY 540 AAACACTTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCAACACAGGAAACGACACA 599
Db 1744 AAACACTTAATAATGGCGAAACGAGTGTGGCGCAAAATCCAGCAACACAGGAAACGACACA 1803
QY 600 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGTTGCAAGCTACTACTAC 659
Db 1804 ATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAACTGTTGCAAGCTACTACTAC 1863
QY 660 GACACCAATCAAGCTAATAATACACCGCAACAACTCAATCAAGCAATACAAATCGCGAGGA 719
Db 1864 GACACCAATCAAGCTAATAATACACCGCAACAACTCAATCAAGCAATACAAATCGCGAGGA 1923
QY 720 ATTAGTGAATCAAAACAGTAATGAAACGACTTTTAAATGATCTACTAAATACAGTATCATCTGT 779
Db 1924 ATTAGTGAATCAAAACAGTAATGAAACGACTTTTAAATGATCTACTAAATACAGTATCATCTGT 1983

780 AAATTCACCTCAAAATTTCTACAAATGCGGAAAATGTTTCAACACGCAAGATACCTTCAAC 839
1984 AAATTCACCTCAAAATTTCTACAAATGCGGAAAATGTTTCAACACGCAAGATACCTTCAAC 2043
840 TGAACCAACACCTTCAAAATGAAATGAGCTCCACAGAGTACAGATGCAAGTAAATGAA 899
2044 TGAAGCAACACCTTCAAAATGAAATGAGCTCCACAGAGTACAGATGCAAGTAAATGAA 2103
900 TGTAGTAAATCAAGCGGTTAATACAAAGTGGCTAGAAATGAGATGAGATTTAGCGGC 959
2104 TGTAGTAAATCAAGCGGTTAATACAAAGTGGCTAGAAATGAGATGAGATTTAGCGGC 2163
960 AGTAGCTGAGATGACCGGAGCTGGACAGATATTACGAATCAGTTGACGAAATGTCAC 1019
2164 AGTAGCTGAGATGACCGGAGCTGGACAGATATTACGAATCAGTTGACGAAATGTCAC 2223
1020 AGTTGGTATTTGACTCTGGTACGACTGCTGATCCGACCAAGCAGGTTATGTCACAACTGAA 1079
2224 AGTTGGTATTTGACTCTGGTACGACTGCTGATCCGACCAAGCAGGTTATGTCACAACTGAA 2283
1080 TTATGGTATTTGACTCTGGTACGACTGCTGATCCGACCAAGCAGGTTATGTCACAACTGAA 1139
2284 TTATGGTATTTGACTCTGGTACGACTGCTGATCCGACCAAGCAGGTTATGTCACAACTGAA 2343
1140 TAAAGAAATTAACCTTAAATGTTGTTAACTTCAAGTGGCTAAAGTGGCTAAAGTGGCTGG 1199
2344 TAAAGAAATTAACCTTAAATGTTGTTAACTTCAAGTGGCTAAAGTGGCTAAAGTGGCTGG 2403
1200 AGATCAAGTATTTGCAATGTTGTTAACTTCAAGTGGCTAAAGTGGCTAAAGTGGCTGG 1259
2404 AGATCAAGTATTTGCAATGTTGTTAACTTCAAGTGGCTAAAGTGGCTAAAGTGGCTGG 2463
1260 AGACTATGTTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1319
2464 AGACTATGTTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2523
1320 CCTGAAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1379
2524 CCTGAAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2583
1380 AGCAAAACAAAACAGTATTTAGTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1439
2584 AGCAAAACAAAACAGTATTTAGTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2643
1440 TAAAGTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1499
2644 TAAAGTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2703
1500 CAATCCAAAGTGGAGATAACGTTTATGCGCGGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1559
2704 CAATCCAAAGTGGAGATAACGTTTATGCGCGGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2763
1560 GGATAGTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1619
2764 GGATAGTAAATGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2823
1620 TGCAGCTGATTTATCTGAAAGTTACTTGTGAAATCCAGAAAATTTGAGGATGTCACATAA 1679
2824 TGCAGCTGATTTATCTGAAAGTTACTTGTGAAATCCAGAAAATTTGAGGATGTCACATAA 2883
1680 TAGTGTGAATTTATCAATTTCCCAATCCCAATCAATTAATTAAGTAGAGTTTAAATACGCTGA 1739
2884 TAGTGTGAATTTATCAATTTCCCAATCCCAATCAATTAATTAAGTAGAGTTTAAATACGCTGA 2943
1740 TGATCAAAATTAACACCGGTATATAGTATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTT 1799
2944 TGATCAAAATTAACACCGGTATATAGTATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTT 3003
1800 AGGTGATTTAGCTTTTACGTTTCACTTTATATGGGTAAATCTCGAATATATTTTGGCGCTC 1859
3004 AGGTGATTTAGCTTTTACGTTTCACTTTATATGGGTAAATCTCGAATATATTTTGGCGCTC 3063
1860 TATGTCATGGGACACGAAGTAGCATTTTAAACCGGATCAGGTTCTGTGTCACGGTATCGA 1919

3064 TATGTCATGGGACACGAAGTAGCATTTTAAATACCGATCAGGTTCTGTGTCACGGTATCGA 3123
1920 TAAACCAAGTTGTTCTCTGAAACCAACCTGATGAGCTTGGTGAATTTGAACCAATTTCCAGAGA 1979
3124 TAAACCAAGTTGTTCTCTGAAACCAACCTGATGAGCTTGGTGAATTTGAACCAATTTCCAGAGA 3183
1980 TTCAGATTCTCACCCAGGTTTCAGATTCTGGCAGCGGATTTCTAATTCAGATAGCGGTTTCA 2039
3184 TTCAGATTCTCACCCAGGTTTCAGATTCTGGCAGCGGATTTCTAATTCAGATAGCGGTTTCA 3243
2040 TTCGGTATGATTTCTACATCAGATGATGTTTTCAGATTTCAGCGAGTATTCAGATTTCAGC 2099
3244 TTCGGTATGATTTCTACATCAGATGATGTTTTCAGATTTCAGCGAGTATTCAGATTTCAGC 3303
2100 AAGTGATTTACAGACTCAGCGAGTATTCAGATTTCAGCGAGGATTTCCGACTCAGCGAGGGA 2159
3304 AAGTGATTTACAGACTCAGCGAGTATTCAGATTTCAGCGAGGATTTCCGACTCAGCGAGGGA 3363
2160 TTTCCGACTCAGACCAATGATTCGGATTCAGATAGCGGATTTCTGACTCAGACAGTACTCAGA 2219
3364 TTTCCGACTCAGACCAATGATTCGGATTCAGATAGCGGATTTCTGACTCAGACAGTACTCAGA 3423
2220 TTTCCGACTCAGACCAATGATTCGGATTCAGATAGCGGATTTCTGACTCAGACAGTACTCAGA 2279
3424 TTTCCGACTCAGACCAATGATTCGGATTCAGATAGCGGATTTCTGACTCAGACAGTACTCAGA 3483
2280 TAGCGATTTACAGATTTCAGATAGCGGATTTCCGACTCAGACAGTACTTCCGACTCAGACAGG 2339
3484 TAGCGATTTACAGATTTCAGATAGCGGATTTCCGACTCAGACAGTACTTCCGACTCAGACAGG 3543
2340 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2399
3544 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 3603
2400 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2459
3604 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 3663
2460 CAGCGATTTACAGATTTCAGATAGCGGATTTCCGACTCAGACAGTACTTCCGACTCAGACAGG 2519
3664 CAGCGATTTACAGATTTCAGATAGCGGATTTCCGACTCAGACAGTACTTCCGACTCAGACAGG 3723
2520 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2579
3724 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 3783
2580 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2639
3784 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 3843
2640 CAGTCACTCAGGATTTCCGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAG 2699
3844 CAGTCACTCAGGATTTCCGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAG 3903
2700 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2759
3904 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 3945
2760 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2819
3946 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 4005
2820 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2879
4006 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 4065
2880 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2939
4066 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 4125
2940 TTTCTGACTCAGACAGTACTTCCGACTCAGACAGCGGATTTCCGACTCAGACAGTACTTCCG 2999

Db 4126 TACTAATGCTTCTTAATAAAAAATGAGGCTAAAGATAGTAAAGAACCACTTACCAGATACAGG 4185
QY 3000 TTCTGAGATGAGCAAAATACGTCACATAATTTGGGATTAATAGCATCAATAGTTCAAT 3059
Db 4186 TTCTGAGATGAGCAAAATACGTCACATAATTTGGGATTAATAGCATCAATAGTTCAAT 4245
QY 3060 ACTACATTTTTCAG 3119
Db 4246 ACTACATTTTTCAG 4305
QY 3120 TTAATCATATGATTCATGAG 3179
Db 4306 TTAATCATATGATTCATGAG 4364
QY 3180 AATATATATGTTTGAAT 3239
Db 4365 AATATATATGTTTGAAT 4424
QY 3240 GTAGATGTTTAT 3299
Db 4425 GTAGATGTTTAT 4484
QY 3300 GGAATAATTAAG 3359
Db 4485 GGAATAATTAAG 4544
QY 3360 TATGCTGATCACAAATTTGGGAAAGTAATCGTCGAGTCAGTGCTTCTGCGGAGAGAG 3419
Db 4545 TATGCTGATCACAAATTTGGGAAAGTAATCGTCGAGTCAGTGCTTCTGCGGAGAGAG 4604
QY 3420 ATCCATATGATCT-AGTCGTTGAACTGACTAATAATAAAAAATAATCTAGAACAGTAG 3479
Db 4605 ATCCATATGATCT-AGTCGTTGAACTGACTAATAATAAAAAATAATCTAGAACAGTAG 4663
QY 3480 AGAGATATAGAGAAAGCTT 3498
Db 4664 AGAGATATAGAGAAAGATT 4682

RESULT 4

US-08-781-986A-587
; Sequence 587, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 587:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-587

Query Match 97.2%; Score 3399.6; DB 4; Length 4709;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3466; Conservative 1; Mismatches 11; Indels 21; Gaps 4;

QY 1 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGTAGGATTAACATTA 60
Db 1204 GGTACCATAAATACACATCTGCTTTTGAATAAATATGATTTCAAGTAGGATTAACATTA 1263
QY 61 GGTAGAGTTTCATATTAATAATAAAAAATGTTTCCAATCAAATCGTAGCTTGTCTTTGTA 120
Db 1264 GGTAGAGTTTCATATTAATAATAAAAAATGTTTCCAATCAAATCGTAGCTTGTCTTTGTA 1323
QY 121 ATTCTTAAATAGCAATAATAATAAATGTTTGTAGTAAAGTATATTTGGATAATAAAA 180
Db 1324 ATTCTTAAATAGCAATAATAATAAATGTTTGTAGTAAAGTATATTTGGATAATAAAA 1383
QY 181 TATCGATACAAAATTAATTTGCTATAATGCAATTTTGTAGTATAATTCATTTAAACAGAGATT 240
Db 1384 TATCGATACAAAATTAATTTGCTATAATGCAATTTTGTAGTATAATTCATTTAAACAGAGATT 1443
QY 241 AATATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTTTAAAAAGAGGGAATAA 299
Db 1444 AATATATATC-TTAAAGGGTATATAGTTAATAATAAATGACTTTTTTAAAAAGAGGGAATAA 1503
QY 300 AATGAATATGAAGAAAAAGAAAAACACCAATTCGAAAAAATCGATTGGCGTGGCTTC 359
Db 1504 AATGAATATGAAGAAAAAGAAAAACACCAATTCGAAAAAATCGATTGGCGTGGCTTC 1563
QY 360 AGTGCTTTGAGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAACAGATGCAAG 419
Db 1564 AGTGCTTTGAGTACGTTAATCGGTTTGGACTACTCAGCAGTAAAGAACAGATGCAAG 1623
QY 420 TGAATAATGTTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAG 479
Db 1624 TGAATAATGTTTACGCAATCTGATAGCGCAAGTAACGAAAGCAAAAGTAATGATTCAAG 1683
QY 480 TAGCGTTAGTCTGCGACCTTAAACACAGACACCAAAAGCTGAGTGATCTAAACATCGTC 539
Db 1684 TAGCGTTAGTCTGCGACCTTAAACACAGACACCAAAAGCTGAGTGATCTAAACATCGTC 1743
QY 540 AAACACTAATAATGCGGAAACAGAGTGTGGCGCAAAATCCAGACAAACAGGAAACGACACA 599
Db 1744 AAACACTAATAATGCGGAAACAGAGTGTGGCGCAAAATCCAGACAAACAGGAAACGACACA 1803
QY 600 ATCATCATCAAAATGCAACTACGGAAGAAACGCGGTAATCTGGTGAAGCTACTACTAC 659
Db 1804 ATCATCATCAAAATGCAACTACGGAAGAAACGCGGTAATCTGGTGAAGCTACTACTAC 1863
QY 660 GACAAAGAAATCAGCTTAATACACCGGCAACACTCAATCAAGCAATACAAATCGGAGGA 719
Db 1864 GACAAAGAAATCAGCTTAATACACCGGCAACACTCAATCAAGCAATACAAATCGGAGGA 1923
QY 720 ATTAGTGAATCAAAACAAAGTAATGAAACGACTTTTAATGATCTAATACAGTATCATCTGT 779
Db 1924 ATTAGTGAATCAAAACAAAGTAATGAAACGACTTTTAATGATCTAATACAGTATCATCTGT 1983
QY 780 AATTCACCTCAAAATTCACAAATCGGAAAAATGTTTCAACAAACGCAAGATACTTCAAC 839
Db 1984 AATTCACCTCAAAATTCACAAATCGGAAAAATGTTTCAACAAACGCAAGATACTTCAAC 2043
QY 840 TGAAGCAACACCTTCAACAAATGAAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGA 899
Db 2044 TGAAGCAACACCTTCAACAAATGAAATCAGCTCCACAGAGTACAGATGCAAGTAATAAGA 2103
QY 900 TGTAGTTAATCAAGCGGTTTAAATACAAAGTGGCGCTAGATGAGAGCATTAGTTTACGGCG 959

; LOCATION: (1)..(5406)
US-09-386-959-1

Query Match	17.7%;	Score 619.2;	DB 4;	Length 5406;
Best Local Similarity	78.8%;	Pred. No. 6.2e-124;		
Matches 738;	Conservative 0;	Mismatches 198;	Indels 0;	Gaps 0;
Qy	1978	GATTCAATGTTCTGACCCAGGTTTCAGATTCTCGGACGCGATTCTAATTCAGATAGCGGTTCA	2037	
Db	4246	GACTCAGACGCAGATAGTGATTCGGATTCGGATAGCGATTTCAGATTCTGATAGTGACTCA	4305	
Qy	2038	GATTCGGGTAGTGNATTCATCATCAGATAGTGGTTTCAGATTTCAGCGAGTGATTTCAGATTCA	2097	
Db	4306	GACTCAGACAGTGTACTCAGATTCCGATAGCGACTTCGGATTTCAGATAGTGATTTCGACGCA	4365	
Qy	2098	GCAAGTGATTTCAGACTCAGCGAGTGATTTCAGATTTCAGCAAGCGATTTCGGACTCAGCGAGC	2157	
Db	4366	GACAGTGACTCAGACTCAGATAGTGACTCGGATTCGGATTCGGATAGTGATTTCGACGACAGCAGC	4425	
Qy	2158	GATTCGCACTCAGCAATGACTCGGATTTCAGATAGCGATTTCGACTCAGACAGTGACTCA	2217	
Db	4426	GATTCGACTCAGATAGTGACTCAGACGCGATAGTGATTTCGGATTTCGGATAGCGATTTCG	4485	
Qy	2218	GATTCGACAGTGACTCAGATTTCAGATAGCGATTTCGACTCAGACAGTGACTCAGATTTC	2277	
Db	4486	GATCGACAGCGACTTCGGATTTCAGATAGTGATTTCGACGCGAGACAGTGACTCAGACTCA	4545	
Qy	2278	GATAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGACAGTGATTTCGACTCAGACAGC	2337	
Db	4546	GATAGTGACTTCGGATTTCGATAGTGATTTCGACGCGACAGACGAGTTCGGATTTCGATAGC	4605	
Qy	2338	GATTCGACTTCGACAGTGTTCGCACTCAGACAGCGATTTCAGATTTCGACAGTGATTTC	2397	
Db	4606	GATTCGACTTCGACAGCGATTTCAGATTTCAGACAGCGACTCAGATTTCGATAGTGATTTC	4665	
Qy	2398	GACTCAGATAGCGATTTCGCACTCAGATAGCGACTCAGATTTCAGACAGCGATTTCAGATTTC	2457	
Db	4666	GATTCAGACAGTGACTTCGGATTTCGATAGTGACTTCAGACTCAGACAGTGACTCAGATTTC	4725	
Qy	2458	GACAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGACAGTGACTCAGATTTCGACAGT	2517	
Db	4726	GATAGCGACTTCAGATTTCAGACAGTGATTTCGACTCAGATAGTGACTTCGGATTTCAGACAGT	4785	
Qy	2518	GACTTCGGATTTCAGATAGCGATTTCAGATTTCGCAAGTGACTCAGATTTCGACAGTGACTCA	2577	
Db	4786	GATTCGGATTTCGATAGCGATTTCGATTCGGATAGTGACTTCGGATTTCAGACAGTGATTTCG	4845	
Qy	2578	GACTTCAGACAGTGATTTCGGATTTCAGCGAGTGATTTCGGATTTCAGATAGTGATTTCGACTCC	2637	
Db	4846	GACTTCAGACAGCGACTTCGGATTTCAGATAGTGATTTCGACTCAGACCGGATTTCGATTTC	4905	
Qy	2638	GACAGTGACTTCGGATTTCAGATAGCGACTCAGACTTCGGATAGCGACTTCGGATTTCAGATAGC	2697	
Db	4906	GATAGTGACTTCGGATTTCAGACAGTGATTTCGCACTTCAGACAGCGACTTCGGATTTCAGATAGT	4965	
Qy	2698	GATTCGAGCTCAGATAGCGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGACAGCGATTTC	2757	
Db	4966	GATTCGCGCGACAGCGACTTCGGATTTCAGATAGTGATTTCGGACGCGACAGCGGATTTC	5025	
Qy	2758	GATTCAGACAGCGACTCAGACAGTGACTTCAGATTTCAGATAGTGACTTCGGATTTCAGCGAGT	2817	
Db	5026	GATAGTGACTTCGGATTTCAGACAGTGATTTCGCACTTCAGACAGCGATTTCGGATTTCAGACAGT	5085	
Qy	2818	GATTCAGACTCAGGTAGTGACTTCGGATTTCATCAGTGATTTCGCACTCAGAAAGTGATTTC	2877	
Db	5086	GACTCGGACTCAGATAGCGACTTCGGATTTCAGACAGTGACTTCGGACTCAGATAGTGACTTC	5145	
Qy	2878	AATAGCGATTTCGGAGTGACTTCCTAACAAATAATGTA	2913	
Db	5146	GATTCAGACAGCGACTTCGGATTTCCTGATAAAAATGCA	5181	

US-09-710-279-3945
; Sequence 3945, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3945
; LENGTH: 4627
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3945

[illegible]

QY 2692 GATAGCGATTTCGAGCTCAGATAGCGATTTCAGATCAGACGCGATTTCAGATCAGACAGC 2751
DB 4174 GATAGCGATTTCGAGCTCAGATAGCGATTTCAGATCAGACGCGATTTCAGATCAGACAGC 4233
QY 2752 GATTCAGATTTCAGACAG-----CGACTCAGACAGTGAATTCAGATTCAGATTCAGTCCG 2805
DB 4234 GACTCCGATTTCAGATAGTTCGAGCTCAGACAGCGATTTCGAGTTCCGATAGTGAATTCG 4293
QY 2806 GATTCAGCGAGTTCAGACTCAGTGTAGTGAATTCAGATTCAGATTCAGATTCAGTCCGACTCA 2865
DB 4294 GATTCAGACAGTTCGAGCTCAGACAGCGATTTCGAGTTCCGATAGTGAATTCGAGTTCCG 4353
QY 2866 GAAAGTGAATTCAGATAGCGATTTCGAGTTCGAGTTCTAAACAATAGTTCGCGCTAAT 2925
DB 4354 GACAGTGAATTCGAGCTCAGACAGCGATTTCGAGTTTCAGACAGTGAATTCGAGTTCAG 4413
QY 2926 TCACCTAAAAATGGTACTAATGCTTCTTAATAAAAAATGAGCTAAGAGATGAAGAACCA 2985
DB 4414 GACTCCGATTTCAGACAGCGACTCGGATTTCGATATAAAAAATGCAAAAGA---TAAATTA 4470
QY 2986 TTACACAGATACAGTTCGAGTGAAGTGAAGCAATACGTCATTAATTTGGGGATTTATTAGCA 3045
DB 4471 GATACAGAGCAAAATGAAGATCATGATTTCTAAAGGCACATTAATTCGAACTTTATTGCA 4530
QY 3046 TCAATAGGTTCTACTACTTTTTCAGAGAAAAAAGAAAAATGAAGATAAGAAATA 3101
DB 4531 GGTTAGGAGCATTAATATTAGGAGCGCTGTAAGAAAGATTAAGAAATA 4586

RESULT 8

US-09-710-279-35

; Sequence 35, Application US/09710279

; Patent No. 6703492

; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS-BLINDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: P03480US

; CURRENT APPLICATION NUMBER: US/09/710,279

; CURRENT FILING DATE: 2000-11-09

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 35

; LENGTH: 1485

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: nucleic acid sequence

US-09-710-279-35

Query Match 16.7%; Score 584.8; DB 4; Length 1485;

Best Local Similarity 73.9%; Pred. No. 9.9e-117;

Matches 774; Conservative 0; Mismatches 262; Indels 12; Gaps 2;

QY 1978 GATTCAGATTTCACCCAGTTTCAGATTTCGGCAGCGATTTCATTAATTCAGATAGCGGTTCA 2037
DB 438 GACTCCGATTTCAGACAGTTCGAGTTCGGACGCGACAGTGAATTCAGATTCAGACAGTTCG 497
QY 2038 GATTCGGGTAGTTCATCATCAGATAGTTCGATTTCAGATTTCAGCGAGTTCAGATTTC 2097
DB 498 GACGCGACAGCGACTTCGGATTTCGATAGTTCGAGCGCAGACAGTTCAGATTTC 557
QY 2098 GCAAGTGAATTCAGACTCAGCGAG-----TGATTCAGATTTCAGCAAGCGATTTCGACTCA 2151
DB 558 GACAGTGAATTCAGACTCAGCGAGTTCGATTTCGATTTCGATTTCAGACAGTTCGATTTC 617
QY 2152 GCGAGCGATTTCGACTCAGCAATGACTTCGATTTCAGATTCGATTTCGATTTCGATTTCG 2211
DB 618 GACAGTGAATTCAGACTCAGCGAGTTCGATTTCGATTTCGATTTCGATTTCGATTTCG 677

RESULT 9

US-09-200-650E-2

; Sequence 2, Application US/09200650E

; Patent No. 6680195

; GENERAL INFORMATION:

; APPLICANT: Patti, Joseph M.

; APPLICANT: Foster, Timothy J.

; APPLICANT: Hook, Magnus A.O.

; APPLICANT: Eidmann, Deirdre Ni

; APPLICANT: Perkins, Samuel L.

; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus

; FILE REFERENCE: P06283US2/BAS

; CURRENT APPLICATION NUMBER: US/09/200,650E

; PRIOR FILING DATE: 1998-11-25

; PRIOR APPLICATION NUMBER: 60/066,815

; PRIOR FILING DATE: 1997-11-26

;; PRIOR APPLICATION NUMBER: 60/098,427
;; PRIOR FILING DATE: 1998-08-31
;; NUMBER OF SEQ ID NOS: 23
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 2969
;; TYPE: DNA
;; ORGANISM: Staphylococcus aureus
US-09-200-650E-2

Query Match 16.3%; Score 568.6; DB 4; Length 2969;
Best Local Similarity 70.3%; Pred. No. 3.9e-113;
Matches 808; Conservative 0; Mismatches 329; Indels 12; Gaps 3;

```
QY 2068 GGTTCAGATTCCAGCGAGTGATTTCAGATTCCAGCAAGTGATTTCAGACTCAGCGAGTGATTCA 2127
DB 1724 GATCCGAACCAAGTCCAGACCCAGACCAGAACCCAGACCCGCGGATCCGGATTTCG 1783
QY 2128 GATTCCAGCAAGCGATTCCGACTCAGCGAGCGATTCCGACTCAGACCAATGACTCCGATTCA 2187
DB 1784 GATTCCAGCAAGCGATTCCGACTCAGCGAGCGATTCCGACTCAGACCAATGACTCCGATTCA 1843
QY 2188 GATAGCGATTTCGACTCAGACAGTGCATCAGATCCGACAGTGCATCAGATTCAGATGAC 2247
DB 1844 GATAGCGATTTCGACTCAGACAGTGCATCAGATCCGACAGTGCATCAGATTCAGATGAC 1903
QY 2248 GATTTCGACTCAGACAGTGCATCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2307
DB 1904 GATTTCGACTCAGACAGTGCATCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1963
QY 2308 GATTCCGACAGTGCATTCGACTCAGACAGCGATTTCGACTCCGACAGTGCATTCGCGACTCA 2367
DB 1964 GATTCCGACAGTGCATTCGACTCAGACAGCGATTTCGACTCCGACAGTGCATTCGCGACTCA 2023
QY 2368 GACAGCGATTTCGACTCCGACAGTGCATTCGACTCCGACAGTGCATTCGCGACTCCGACTCAG 2427
DB 2024 GATAGCGACTCAGACAGTGCATTCGACTCCGACAGTGCATTCGCGACTCCGACTCCGACTCAG 2083
QY 2428 GACTCAGATTTCAGACAGCGGATTTCAGATTTCAGACAGCGGATTTCAGATTTCAGATGCGATTCA 2487
DB 2084 GATTTCAGACTCAGATAGCGGATTTCAGACTCAGATAGCGGATTTCAGATTTCAGACAGCGGATTCA 2143
QY 2488 GATTCCGACAGTGCATTCGACTCCGACAGTGCATTCGACTCCGACAGTGCATTCGCGACTCC 2547
DB 2144 GATTCCGACAGTGCATTCGACTCCGACAGTGCATTCGACTCCGACAGTGCATTCGCGACTCC 2203
QY 2548 GACAGTGCATTCGACTCCGACAGTGCATTCGACTCCGACAGTGCATTCGCGACTCCGACTCAG 2607
DB 2204 GATAGCGGATTTCAGACTCAGACAGCGGATTTCAGACTCAGACAGCGGATTTCAGACTCAGATG 2263
QY 2608 GATTCCGATTTCAGATAGTGCATTCCGACTCCGACAGTGCATTCCGATTTCAGATAGCGGATTCA 2667
DB 2264 GACTCAGATTTCAGATAGGCGATTTCAGACTCAGACAGCGGATTTCAGATTTCAGATAGCGGATTTCG 2323
QY 2668 GACTCCGATAGCGGATTTCGATTTCAGATAGGCGATTTCGACTCAGACAGCGGATTTCAGATTTCAGATG 2727
DB 2324 GACTCAGACAGCGGATTTCAGATTTCAGACAGCGGATTTCAGACTCCGATAGCGGATTTCAGATTTCAG 2383
QY 2728 GACAGCGATTTCAGATTCAGACAGCGGATTTCAGATTTCAGACAGCGGATTTCAGACTCAGGTTGACTGCC 2781
DB 2384 GATAGCGGATTTCGATTTCAGACAGTGCATTTCAGATTTCAGACAGCGGATTTCAGACTCCGATAGCG 2443
QY 2782 GACTCAGATTTCAGATAGTGCATTCCGATTTCAGCGGATTTCAGACTCAGGTTGACTGCCGTTGACTGCC 2841
DB 2444 GACTCAGACTCAGACAGCGGATTTCAGACTCAGATAGCGGATTTCAGACTCCGATAGCGGATTTCG 2503
QY 2842 GATTTCAGAGTGCATTCCGACTCAGAAAGTGCATTTCAGAAATAGCGGATTTCGAGTCCGTTCT 2901
DB 2504 GATTTCAGATAGCGGATTTCAGACTCAGATAGTGCATTTCGAGTCCGATTTCAGAGGTTTACACCAAT 2563
QY 2902 AACCAATAGTGCATTCCGCTTAATTCACCTAAAAATGCTACTTAATGCTTCTTAATAAAT 2961
DB 2564 AATGAACAGAAAGCCCAATCAAT---CCTAAGGTGAAGTAACCAATTCCTAATAAGGTA 2620
```

```
QY 2962 GAGGCTAAGATAGTAAAGAACCACTTACCAGATACAGGTTCTTGA---AGATGAAGCAAT 3018
DB 2621 TCAGAAACACACAAAACGATGCTTTTACCAGAACAGGAGATAGAGCGAAACACAAAT 2680
QY 3019 ACCTCAGTAATTTGGGATTTATAGCATCAATAGGTTCACTACTACTTTTTCAGAGAAAA 3078
DB 2681 GCAACTTTATTTGGTCAATGATGCGATTATAGGATCATTACTATTGTTTAGAAAAACGC 2740
QY 3079 AAAGAAATAGATAGAAATAGTAATATATATATTAATTAATATATATATATATATATAT 3138
DB 2741 AAGCAAGATCATAAGAAAAAGCGTAAATACTTTTGTAGCGCGAATACATTTGTTATTCGG 2800
QY 3139 AGAAGCCACCTTAAAGGTGCTTCTTTTACTTGGATTTCCTCAAAATATATATTTGTTGAATAT 3198
DB 2801 TTTTGTGTTGAAATGATTTTAAAGTGAATGATTAGCGTAAATGTTGATTAAGTAG 2860
QY 3199 AATTAATAA 3207
DB 2861 AATTAGAAA 2869
```

RESULT 10

US-08-956-171E-96
; Sequence 96, Application US/08956171E
; Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956.171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB24BP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:

LENGTH: 11050 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 96:

US-08-956-171E-96

Query Match

14.6%; Score 510; DB 4; Length 11050;

[illegible]

2598	Db	TACTATCCAGATCCAAATGACTCTAACCTTTAAAGAAATAACAGACCCTAAATTTTAAAAATAGA	2657
1702	Qy	AATCCAAATCAATATATAAGTAGAGTTTTAATACCGCTGATGATCAAAATTTCAACACCGGTAT	1761
2658	Db	ATCTATTATGAGCAICCAAAATGTAGCTAGTATTAAATTTGGTGATATTCTTAAACATAT	2717
1762	Qy	ATAGTAGTTGTTTAATGTCATATTGTATCCGAAATAGCAAAAGTGATTTAGCTTTTAGCTTCA	1821
2718	Db	GTAGTATTATAGAAAGGCGCATACGA-CAATAACAGGTAAAGAACTTAAAAACTCAGAGTTAT	2776
1822	Qy	ACTTTATATCGGTATAACTCGAATATATTTGGCGCTCTATGTCATCGGACAACGAGTA	1881
2777	Db	TCAGAAAAATGTTGATCCTCTGACAAATAGAGACTACAGTATTTTCGGTTGGAAATATAGA	2836
1882	Qy	GCATTTTAATAACGGATCAGGTTCTGGTGTACGGGTATTCGATAAAACAGTGTGTTCTCTGAACAA	1941
2837	Db	GAATGTGTAC-----GTTATGGTGGTGAAGTCTCTGATGGTGATTCAGCAGTAAAT	2888
1942	Qy	CTTGATGAGCCTGGTGAAATTGAACCAATTTCCAGAGGATTCAGATTCAGACCCAGGTTCA	2001
2889	Db	CCGAAAGACCCAACTCCAGGCGCGCGGTTGACCCAGAAACCAAGTCCAGACCCAGAACCA	2948
2002	Qy	GATTCCTGGCAGCATCTTAATTCAGATAGCGGTTTCAGATTCGGGTAGTGATTTCTACATCA	2061
2949	Db	GAACCAACGCCGATCCAGAACCAAGTCCAGACCCGAAACCGGAAACCAAGCCACAGACCGG	3008
2062	Qy	GATAGTGGTTTCAGATTCAGCGAGTGATTCAGATTCAGCAAGTGATTCAGACTCAGCGAGT	2121
3009	Db	GATCCGGATTCGGATTCAGACAGTGACTCAGGCTCAGACGCGACTCAGGTTTCAGATAGC	3068
2122	Qy	GATTCAGATTCAGACAGGATTCGGACTCAGCGAGCGGATTCGGACTCAGACAATGACTCG	2181
3069	Db	GACTCGAAATCAGATAGCGATTCGGATTCAGACAGTGATTCAGATTCAGACAGCGACTCA	3128
2182	Qy	GATTCAGATAGCGATTCGACTCAGACAGTGACTCAGATTCCGACAGTGACTCAGATTCA	2241
3129	Db	GAATCAGATAGCGACTCAGAAATCAGATGAGTGAGTTCAGATTCAGACAGTGACTCGGACTCA	3188
2242	Qy	GATAGCGATTCGACTCAGACAGTGACTCAGATTCAGATAGCGGATTCAGATTCAGATAGC	2301
3189	Db	GACAGTGATTCAGACTCAGATAGCGGATTCAGACTCAGATAGCGGATTCAGACTCAGACAGC	3248
2302	Qy	GATTCAGATTCGACAGAGTGATTCGGACTCAGACAGCGGATTCGACTCCGACAGTGATTC	2361
3249	Db	GATTCAGATTCAGACAGCGACTCAGATTCAGACAGCGGACTCAGACTCAGATAGCGACTCA	3308
2362	Qy	GACTCAGACAGCGATTCAGATTCGGACAGTGATTCGACTCAGATAGCGGATTCGGACTCA	2421
3309	Db	GACTCAGACAGCGACTCAGATTCAGATAGCGGATTCAGACTCAGACAGCGACTCAGACTCA	3368
2422	Qy	GATAGCGACTCAGATTCAGACAGCGGATTCAGATTCAGACAGCGGATTCAGATTCAGATAGC	2481
3369	Db	GACAGCGACTCAGACTCAGATAGCGGACTCAGATTCAGATAGCGGATTCAGACTCAGACAGC	3428
2482	Qy	GATTCAGATTCGACACAGTGACTCAGATTCGACAGTGACTTCGGATTCAGATAGCGGATTC	2541
3429	Db	GACTCAGATTCAGATAGCGGATTCGGACTCAGACAGCGGATTCAGATTCAGACAGCGACTCA	3488
2542	Qy	GATTCGACAGTGACTCAGATTCGGACAGTGACTCAGACTCAGACAGTGATTCGGATTC	2601
3489	Db	GACTCGGATAGCGATTCAGATTCAGATAGCGGATTCGGGATTCAGACAGTGATTCAGATTC	3548
2602	Qy	GCAGGTGATTCGGATTCAGATAGTGATTCGCGCTCCGACAGTGACTCTCGGATTCAGATAGC	2661
3549	Db	GACAGCGACTCAGACTCGGATAGCGACTCAGATTCAGACAGCGGATTCAGACTCAGATAGC	3608
2662	Qy	GACTCAGACTTCGGATAGCGACTCGGATTCAGATAGCGGATTCGGACTCAGATAGCGGATTC	2721
3609	Db	GACTCAGACTTCGGATAGCGACTCGGATTCAGATAGCGGACTCAGACTCAGATAGTGACTCC	3668
2722	Qy	GAATCA	2727
3669	Db	GAATCA	3674

RESULT 11

US-08-781-986A-96
; Sequence 96, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11050 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-96

Query Match 14.6%; Score 510; DB 4; Length 11050;
Best Local Similarity 55.2%; Pred. No. 2.7e-100;
Matches 1162; Conservative 0; Mismatches 905; Indels 39; Gaps 7;
QY 652 ACTACTAGCAACGAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAAT 711
DB 1578 AATACCCTCAACACAGAGCCAGCTTCAACAAATGAACACCTCAACCCGCGCAATTAA 1637
QY 712 GCGGAGGAATAGTGAATCAACAGTAATGAACGACTTTAATGATCTAATACAGTA 771
DB 1638 AATCAAGCAACTGTGCAAAATCAAGATCAAACTGTTCTCAAGAGCAAAATTCCTCA 1697
QY 772 TCATCTGTAATTCACCTCAAAATTCCTCAAAATGCGGAAATGTTTCAACCAACGCAAGT 831
DB 1698 GTAGATATAAACAACCAAGTATGCTAATAGCATAGCAACAAACAGTGAGCTTAAAT 1757
QY 832 ACTTCAACTGAAGCAACACCTTCAACAAATGAATCAAGTCCACAGAGTACAGATGCAAGT 891
DB 1758 TCTCAACATTTAGATTTTACCACATCATCAACCAACAAACGATTTCCAAATGCGCAAGAACT 1817
QY 892 AATAAGATGTAGTTAATCAAGCGTTAATCAAGTGGCGCTAGAAATGAGAGCAATTAGT 951
DB 1818 AGTAAACCAAGTTTAGAACGAGAGCTGTAAGTAGTTTAGCTTGTGCTGAACCGGTAGTA 1877
QY 952 TTAGCGGAGTAGCTGCAGATGCACCGGAGCTGGCAAGATATTTACGAATCAGTTGACG 1011
DB 1878 AATGCTGCTGATGCTAAAGGTACAATGTAATGATGAAGTTACGGCAAGTAATTTCAAG 1937
QY 1012 AATGTACAGTTGGTATTGACTCTCTGTAGACTGTGTATCCGCCAACCAAGAGTTATGTC 1071

DB 1938 TTAGAAAAGAGCTACATTTGACCCCTAAATCAAAGTGGTAACACATTTATGGCGCAAAATTTT 1997
QY 1072 AAACGTGAATTTATGGTTTTTCAGTGCCTAAATTTCTGCTGTTAAAGGTGACACATTCAAATA 1131
DB 1998 ACAGTGACAGATAAAGTGAATCAGGGGATTTATTTACAGCGAAGTTTACAGATAGTTTA 2057
QY 1132 ACTGTACCTAAAGAAATTAACCTTAAATGGTGTAACTTCAACTGTCTAAAGTGCACCAATT 1191
DB 2058 ACTGGTAATCGAGACGTGGATTTATTTCAATTTCAATATACGATGCCAATTCGACACATT 2117
QY 1192 ATGGCTGGAGATCAAGTATTCGGAAAATGGGTAAATCGATAGTGGTAAATGTTATTTAT 1251
DB 2118 AAAAGTACGAATGGCGATGTTGTAGCTAAAGCAACATATGATATCTTGACTAAGACGAT 2177
QY 1252 ACA-----TTTACAGACTATGTAATACTAAAGATGATGTAAGAGCAACTTTGACCATG 1305
DB 2178 ACATTTGCTTTTACAGATTATGTAATTAATAAGAAAATATTTAAACGCAATTTTCATTA 2237
QY 1306 CCCGCTTATATTGACCTTGAAAATTTTAAAGACACAGTAAATGTGACATTTGGCTACTGCG 1365
DB 2238 CCTTTATTTACAGACCGAGCAAGGCCTTAATCAGGAACATATGATGCGAATTAAT 2297
QY 1366 ATAGTAGTACAACAGCAAAACAAACA-----GTATTAGTAGATTATGAAAATAATGTA 1420
DB 2298 ATTGGGATGAATGTTTAAATAAATAATTAATACTATAGTTCCGCAATTCAGGA 2357
QY 1421 AGTTTTA-----TAACTTATCTATTAAAGGTACAATTTGACCAATCGATAA 1467
DB 2358 ATTGATAAACCAATGGCGCAACATTTCTTCTCAAAATTTGGTGTAGATACAGCTTCA 2417
QY 1468 ACAATAATACGTATCGTCAGACAAATTTATGTCAATCCAA---GTGGAGATAAAGTTATT 1524
DB 2418 GGTCAAAACACATACAGCAACAGTATTTCTTAACCTTAAGCAACGAGTTTAGGTAA 2477
QY 1525 GCGCGGCT---TTTAAACAGGTAATTTAAACCAATACGAGTAGTAAATGCAATTAATAGAT 1581
DB 2478 ACGTGGGTGTATATTAAAGGCTACCAAGATAAAATCGAAGAAAGTAGCGGTAAAGTAA 2537
QY 1582 CAGCAAAATACAAGTATTAAAGTATATAAAGTATATAAGTATGAGCTGATTTATCTGAAAGT 1641
DB 2538 GCTACAGATCAAAACCTGAGAAATTTTGAAGTGAATGATACATCTAAATTTACAGATAGC 2597
QY 1642 TACTTTGTGAATCCAGAAAACCTTTGAGGATGCTCAATATAGTGTGAATTAATACATCCCA 1701
DB 2598 TACTATGAGATCCAAATGACTCTTAACCTTAAAGAAAGTAAACAGACCAATTTAAATAATGA 2657
QY 1702 AATCCAAATCAATATAAGTAGAGTTTAAATAGCCTGATGATCAAAATTTACAAACCCGTAT 1761
DB 2658 ATCTATTATGAGCATCCAAATGTAGTATTAAATTTGGTGATTTACTTAAACATAT 2717
QY 1762 ATAGTAGTTGTTAATGGTCAATATTGATCCGAATAGCAAGGTGATTTAGCTTTAGCTTCA 1821
DB 2718 GTAGTATTAGTAGAGGCAATTACGA-CAATACAGTGAAGAACTTAAAACTCAGGTTAT 2776
QY 1822 ACTTTATATGGGTATACTCGAATATAATTTGGCGCTCTATGTCTATGGGACCAACGAAGTA 1881
DB 2777 TCAAGAAAATCTTGATCCTGTAAACAAATAGAGACTACAGTATTTTCGGTTTGAATAATGA 2836
QY 1882 GCATTTAATACGATCAGGTTTCTGGTGACCGGTATCGATAAACAGGTTGTTCTCTGAACAA 1941
DB 2837 GAATGTTGTAC-----GTTATGGTGGTGAAGTCTGATGGTGAATTCAGCAGTAAAT 2888
QY 1942 CCTGATGAGCTGGTGAATTTGAACCAATTCAGAGGATTCAGATTTCTGACCCAGGTTCA 2001
DB 2889 CCGAAGACCAACTCTCAGGCGCCCGGTTGACCCAGNACCAAGTCCAGACCCAGAACCA 2948
QY 2002 GATTCTGCAGGATTTAAATTCAGATAGCGGTTTCAAGTTCCGTTAGTGAATTTACATCA 2061
DB 2949 GAAACCAAGCCAGATCCAGAACCAAGTCCAGACCCAGAACCCGGAACCAAGCCAGACCG 3008
QY 2062 GATAGTGGTTTCAAGATTCAGCGAGTGAATTCAGATTCAGCAAGTGAATTCAGATCTCAGCGAGT 2121

:	NAME/KEY:	CDS
:	LOCATION:	(33) .. (3308)
US-09-147-405B-14		
	Query Match	13.2%; Score 462; DB 4; Length 3600;
	Best Local Similarity	75.6%; Pred. No. 3.9e-90;
	Matches 573; Conservative	0; Mismatches 185; Indels 0; Gaps 0;
Qy	1984	GATTTGACCCAGGTTCAAGATTCTGGCAGCGATTCTAAATTCAGATAGCGGTTCAGATTGC 2043
Db	2412	GATGACGAACAGGATGCTGATGGGAGAAGAGTTCATGTAAACAATTACTGATCATGATGAC 2471
Qy	2044	GGTAGTGATTCTACATCAGATAGTGGTTCCAGATTCCAGCGAGTGATTCCAGATTCCAGCAAGT 2103
Db	2472	TTTTAGTATAGATAACGGATACCTAATGATGACGAATCCGATTCGGATAGTGACTCAGACAGC 2531
Qy	2104	GATTCAGACTCACGCGAGTGATTCAGATTCCAGCAAGCGAATTCGGAATCCAGCGAGCGATTCC 2163
Db	2532	GACTCAGATCCGATAGTAGTATTCCAGATCCGATAGCGACTCGGATTCAGACACGCCACTCA 2591
Qy	2164	GACTCAGCAATGACTCCGGATTCAGATAGCGATTCTGACTCCAGACAGTGACTCAGATTCC 2223
Db	2592	GATTCAGACAGCGACTCGGATTCGTATAGCGACTCCGATTCAGACACGGACTCAGACTCA 2651
Qy	2224	GACAGTGACTCAGATTCCAGATAGCGATTCTGACTCCAGACAGTGACTCAGATTCCAGATAGC 2283
Db	2652	GACAGTGATTCAGATTCCAGACCGGACTCAGATTCCGATAGTGATTCCAGACTCAGACAGC 2711
Qy	2284	GATTCAGATTCAGATAGCGATTCCAGATTCGCGACAGTGATTCGGAATCCAGACAGCGATTCT 2343
Db	2712	GACTCAGATTCTGATAGTGATTCCAGATCCAGACAGTGATTCAGATTCCAGACAGCGACTCA 2771
Qy	2344	GACTCCGACAGTGATTCCCGACTCAGACAGCGATTCCAGATTCGACAGTGGATTCCGACTCA 2403
Db	2772	GATTCGGATAGTGATTCCAGACTCCAGACAGCGACTCAGATTCCGATAGTGATTCCAGACTCA 2831
Qy	2404	GATAGCGATTCCGACTCAGATAGCGACTCAGATTCCAGACAGCGATTCCAGATTCAGACAGC 2463
Db	2832	GACAGCGACTCAGATTCTGATAGTGATTCCAGATCCAGACAGTGATTCCAGACTCAGACAGT 2891
Qy	2464	GATTCAGATTCCAGATAGCGATTCCAGATTCGCGACAGTGACTCAGATTCGACAGTGACTCG 2523
Db	2892	GATTCAGATTCCGATAGTGATTCCAGATCCGATAGCGACTCCAGACTCGGATAGTGACTCA 2951
Qy	2524	GATTCAGATAGCGATTCCAGATTCGCGACAGTGACTCAGATTCCGACAGTGACTCAGACTCA 2583
Db	2952	GATTCGATAGTGATTCCAGACTCCGATAGCGACTCAGACTCCGATAGTGACTCAGATTCT 3011
Qy	2584	GACAGTGATTCCGATTCCAGCGAGTGATTCCGATTCCAGATAGTGATTCCGACTCCGACAGT 2643
Db	3012	GATAGTGATTCCAGACTCAGACAGCGACTCAGATTCTGATAGTGATTCCAGACTCCAGTCA 3071
Qy	2644	GACTCCGATTCCAGATAGCGACTCAGATCCGATAGCGACTCCGATTCCAGATAGCGATTTCG 2703
Db	3072	GATTCAGATTCCGATAGTGATTCCAGACTCAGGACAGTGATTCCGATTCCGATAGTGATTCA 3131
Qy	2704	GACTCCAGATAGCGATTCCAGAAATCAGACAGCGATTCCAGA 2741
Db	3132	GACTCAGCAACGACTCAGATTTAGGCAATAGCTCAGA 3169

```

RESULT 13
US-08-956-171E-454/c
; Sequence 454, Application US/08956171E
; Patent NO. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequence

```

```

;
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 454:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 454:
US-08-956-171E-454

Query Match 12.8%; Score 448.4; DB 4; Length 4358;
Best Local Similarity 71.4%; Pred. No. 3.5e-87;
Matches 641; Conservative 0; Mismatches 236; Indels 21; Gaps 3;

Qy 1865 CATGGGACACGAAGTAGCATTTAATACGGATCAGGTTCTGGTACCGGTATCGATAAAC 1924
Db |||||
Qy 1925 CAGTTGTTCTGCACACCTCATGAGCTGTGTAATGAACCAATTCAGAGGATTCAG 1984
Db |||||
Qy 2314 AAGTTATCTTTGAAAAACCTGCTGCTTTAACTCAACAGGTACAAATACAACTGAAGATG 2255
Db |||||
Qy 1985 ATTCTGACCCAGGTTTCAGATTCTGCGACGCGATTCTTAATTCAGATAGCGGTTTCAGATTGG 2044
Db |||||
Qy 2254 ATAAGATGCGAGTGGTGGCGAAGTTGATGTAACAATTCAGGATCATGAT---GATTCA 2198
Db |||||
Qy 2045 GTAGTGATTTACATCAGATAGTGGTTTCAGATTTCAGGAGTGTATTCAGATTTCAGCAAGTG 2104
Db |||||
Qy 2197 CACTTGATTAATGGCTACTACGAGAAGAAACATCAGATAGGAGCTCAGATTCTGCACGCG 2138
Db |||||
Qy 2105 ATTTCAGATCAGCGATGATTCAGATTTCAGCAAGGATTCGACTTCAGCGAGCGATTCCG 2164
Db |||||
Qy 2137 ATTTCAGATCAGATAGGAGCTCAGATTTCAGATAGCGACTCAGATTTCAGACAGCGAT---- 2082
Db |||||
Qy 2165 ACTCAGACATGACTCGGATTCAGATAGCGATTCTGACTCAGACAGTGTACTCAGATTCCG 2224
Db |||||
Qy 2081 --TCAGACAGCGACTCAGACTCAGATAGCGATTTCAGATTTCAGACGCGACTCAGACTCAG 2024
Db |||||
Qy 2225 ACAGTGACTCAGATTCAGATAGCGATTCTGACTCAGACAGTGTACTCAGATTTCAGATAGCG 2284
Db |||||
Qy 2023 ACAGCGATTCAGATTCGATAGCGACTCAGACTCAGATAGGAGCTCAGATTTCGATAGCG 1964
Db |||||
Qy 2285 ATTTCAGATTTCAGATAGCGATTTCAGATTTCGACAGTGTATTCGACTTCAGACAGCGATTCTG 2344
Db |||||
Qy 1963 ACTCAGATCAGATAGCGATTTCAGATTTCAGATAGCGATTTCGATTCAGACAGTGTATTC-- 1906
Db |||||

;
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 454:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 454:
US-08-956-171E-454

Query Match 12.8%; Score 448.4; DB 4; Length 4358;
Best Local Similarity 71.4%; Pred. No. 3.5e-87;
Matches 641; Conservative 0; Mismatches 236; Indels 21; Gaps 3;

Qy 2345 ACTCGACAGTGATTCGACTCAGACAGCGATTTCAGATTTCGACAGTGTATTCGACTCAG 2404
Db |||||
Qy 1905 -----AGATTCAGACTCAGATAGCGACTCAGATTCTGACAGCGATTTCAGACTCAG 1856
Db |||||
Qy 2405 ATAGCGATTCGACTCAGATAGCGACTCAGATTTCAGACAGCGATTTCAGATTTCAGACGCG 2464
Db |||||
Qy 1855 ACAGCGACTCAGACTCAGACAGTGTATTCAGATTTCAGACAGCGACTCAGATTTCAGATAGCG 1796
Db |||||
Qy 2465 ATTTCAGATTTCAGATAGCGATTTCAGATTTCGACAGTGTACTCAGATTTCGACAGTGTACTCGG 2524
Db |||||
Qy 1795 ACTCAGACTCAGATAGCGACTCAGATTTCAGATAGCGATTTCGAGTTCGACACAGCGACTCAG 1736
Db |||||
Qy 2525 ATTTCAGATAGCGATTTCAGATTTCGACAGTGTACTCAGATTTCGACAGTGTACTCAGACTCAG 2584
Db |||||
Qy 1735 ATTTCAGATAGCGATTTCAGATTTCAGATAGCGACTCAGATTTCGACAGCGATTTCAGACTCAG 1676
Db |||||
Qy 2585 ACAGTGATTCGATTTCAGCGAGTGTATTCGAGTGTATTCGAGTGTATTCGAGTGTATTCGAGTGT 2644
Db |||||
Qy 1675 ATAGCGATTTCAGACTCAGACAGCGATTTCAGATTTCAGATAGCGACTCAGACTCAGATAGCG 1616
Db |||||
Qy 2645 ACTCGATTTCAGATAGCGACTCAGACTCAGATAGCGACTCAGATTTCAGATAGCGATTTCG 2704
Db |||||
Qy 1615 ACTCAGACTCGGATAGCGATTTCAGATTTCAGACAGCGACTCAGATTTCAGATAGCGATTTCG 1556
Db |||||
Qy 2705 ACTCAGATAGCGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTC 2762
Db |||||
Qy 1555 ACTCAGACAGCGACTCAGATTTCAGATAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTC 1498
Db |||||

;
; RESULT 14
; US-08-781-986A-454/c
; Sequence 454, Application US/08/781986A
; Patent No. 6737248
;
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 454:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-454
;
; Query Match 12.8%; Score 448.4; DB 4; Length 4358;

```

```

Best Local Similarity 71.4%; Pred. No. 3.5e-87;
Matches 641; Conservative 0; Mismatches 236; Indels 21; Gaps 3;

QY 1865 CATGGGACACGAGTAGCATTTAATAACGAGTACAGGTCTCGTGAGCGGTATCGATAAAC 1924
    |||
Db 2374 CAACTGARACAGATAAAATGGTAAATACCGCTTTTGATATTTAGATAGTGGTAAATACA 2315

QY 1925 CAGTTGTTCTGAACACCTGATGAGCCTGGTGAATTAACCAATTCAGAGGATTCAG 1984
    |||
Db 2314 AGTTATCTTTGAAACCTGCTGGCTTAACCAACAGGTACAAATACAACTGAAGATG 2255

QY 1985 ATTCTGACCCAGGTTTCAGATTTCTGGCAGCGATTTCTAATTCAGATAGCGGTTTCAGATTCGG 2044
    |||
Db 2254 ATAAAGATCCGATGGTGGCGAGTTGATGAACAAATACGGATCAATGAT---GATTTCA 2198

QY 2045 GTAGTGAATCTACATCAGATAGTGGTTTCAGATTCAGCGAGTGAATTCAGATTCAGCAAGTG 2104
    |||
Db 2197 CACTTGATAATGCTACTACGAAGAAGAAACATCAGATAGCGACTCAGATTTCTGACAGCG 2138

QY 2105 ATTCAGACTCAGCGAGTTCAGATTCAGATTCAGCAAGCGATTCGACTCAGCGAGCGATTCGG 2164
    |||
Db 2137 ATTCAGACTCAGATAGCGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2082

QY 2165 ACTCAGACAAATGACTCGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2224
    |||
Db 2081 --TCAGACAGCGACTCAGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2024

QY 2225 ACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2284
    |||
Db 2023 ACAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1964

QY 2285 ATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2344
    |||
Db 1963 ACTCAGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1906

QY 2345 ACTCCGACAGTGTTCGACTCAGACAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2404
    |||
Db 1905 -----AGATTCAGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1856

QY 2405 ATAGCGATTCGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2464
    |||
Db 1855 ACAGCGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1796

QY 2465 ATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2524
    |||
Db 1795 ACTCAGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1736

QY 2525 ATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2584
    |||
Db 1735 ATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1676

QY 2585 ACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2644
    |||
Db 1675 ATAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1616

QY 2645 ACTCCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2704
    |||
Db 1615 ACTCAGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1556

QY 2705 ACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2762
    |||
Db 1555 ACTCAGACACGACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1498

```

```

RESULT 15
US-09-710-279-31/c
; Sequence 31, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279

```

```

; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-31

Query Match      12.5%; Score 436.8; DB 4; Length 756;
Best Local Similarity 78.5%; Pred. No. 5.9e-85;
Matches 567; Conservative 0; Mismatches 137; Indels 18; Gaps 3;

QY 2074 GATTCAGCGAGTGATTCAGATTCAGCAAGTGTGATTCAGACTCAGCGAGTGTGATTCAGATTC 2133
    |||
Db 727 GATTCAGACAGTGAATTCAGACGAGATAGTGTGATTCAGACTCAGCGAGTGTGATTCAGCGCA 668

QY 2134 GCAAGCGATTCGAGTTCAGCGAGCGATTCGAGTCCAGACTCAGCAATGACTCGGATTCAGATAGC 2193
    |||
Db 667 GACAGCGATTCGAGCGACAGCACTAACTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 608

QY 2194 GATTCGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2247
    |||
Db 607 GACTCGGATTCGATAGTGTGATTCGAGTTCGATTCAGACAGTGAATTCAGATTCAGATTCAGATTCAG 548

QY 2248 GATTCGACTCAGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2307
    |||
Db 547 GATTCGACGACAGTGAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 488

QY 2308 GATTCGCGAGT-----GATTCGCGACTCAGACAGCGATTCGAGTCCGACAGTGAATTC 2361
    |||
Db 487 GATTCGATAGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCG 428

QY 2362 GACTCAGACAGCGATTCAGATTCGCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 2421
    |||
Db 427 GACGCGACAGTGAATTCAGACGCGATAGTGTGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 368

QY 2422 GATAGCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGATTCAG 2481
    |||
Db 367 GATAGTGAATTCGAGTTCGAGCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 308

QY 2482 GATTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 2541
    |||
Db 307 GACTCGGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 248

QY 2542 GATTCGCGAGTGAATTCAGATTCGCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 2601
    |||
Db 247 GACGCGACAGCGACTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 188

QY 2602 GCGAGTGAATTCGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2655
    |||
Db 187 GACAGTGAATTCGAGTTCAGACAGCGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAGATTCAG 128

QY 2656 GATAGCGACTCAGATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 2715
    |||
Db 127 GACAGTGAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCAG 68

QY 2716 GATTCAGATTCAGACAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2775
    |||
Db 67 GATTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 8

QY 2776 GA 2777
    |||
Db 7 GA 6

```

```

Search completed: November 5, 2004, 12:59:33
Job time : 285 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 06:05:04 ; Search time 117 Seconds
(without alignments)
6924.920 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaccataattacacatc.....gaagagtataagaagaagctt 3498

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 693957 seqs, 115811272 residues

Total number of hits satisfying chosen parameters: 1387914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1592.6	45.5	1651	8	US-60-615-573-18588
2	1393.6	39.8	1600	8	US-60-615-573-18330
3	217.8	6.2	404	8	US-60-615-573-14144
4	169.4	4.8	180	8	US-60-615-573-18462
5	141.4	4.0	4050	6	US-10-556-157-5134
6	99	2.8	171	8	US-60-615-573-14132
7	85	2.4	98412	6	US-10-746-294A-30
8	79.8	2.3	2701	8	US-60-615-573-18342
9	79.4	2.3	388	6	US-10-220-366A-1506
10	73.8	2.1	427	6	US-10-220-366A-1495
11	72.6	2.1	2493	8	US-60-615-573-14147
12	71.6	2.0	5192	8	US-60-615-573-292-360
13	71.2	2.0	404	8	US-60-615-573-14144
14	67.2	1.9	105306	6	US-10-746-294A-111
15	67	1.9	2022	8	US-60-615-573-18593
16	66.8	1.9	337	8	US-60-615-573-14146
17	65.6	1.9	5469	6	US-10-556-157-1983
18	64.4	1.8	1601	8	US-60-615-573-18596
19	64	1.8	47475	6	US-10-915-740A-13
20	64	1.8	2242716	6	US-10-915-740A-1068
21	61.8	1.8	74700	1	PCT-US03-40884-3
22	60.4	1.7	193	8	US-60-615-573-17864
23	60	1.7	129757	6	US-10-746-294A-88
24	57	1.6	384	6	US-10-220-366A-2561
25	55.2	1.6	863	6	US-10-479-081-32

Sequence 61, Appl
Sequence 1, Appl
Sequence 69, Appl
Sequence 5849, Ap
Sequence 18, Appl
Sequence 18214, A
Sequence 60, Appl
Sequence 14092, A
Sequence 3, Appl
Sequence 10577, A
Sequence 16, Appl
Sequence 18495, A
Sequence 78, Appl
Sequence 80, Appl
Sequence 3, Appl
Sequence 4179, Ap
Sequence 59, Appl
Sequence 65, Appl
Sequence 121, App

ALIGNMENTS

RESULT 1
US-60-615-573-18588

; Sequence 18588, Application US/60615573

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William M

; APPLICANT: Murphy, Ellen

; APPLICANT: Olmsted, Stephen

; FILE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES

; TITLE REFERENCE: 031896-084099 (AM 101724)

; CURRENT APPLICATION NUMBER: US/60/615,573

; CURRENT FILING DATE: 2004-10-05

; NUMBER OF SEQ ID NOS: 18598

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 18588

; LENGTH: 1651

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-60-615-573-18588

Query Match 45.5%; Score 1592.6; DB 8; Length 1651;

Best Local Similarity 93.2%; Pred. No. 0;

Matches 1538; Conservative 111; Mismatches 1; Indels 1; Gaps 1;

QY	301	ATGAATATGAGGAAAGAAAGAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA	360
DB	1	ATGAATATGAGGAAAGAAAGAAACACGCAATTCGGAAAAAATCGATTGGCGTGGCTTCA	60
QY	361	GTGCTTGTAGGTACGTTAATCGTTTGGACTACTCAGCAGTAAAGACAGATGCAAGT	420
DB	61	GTGCTTGTAGGTACGTTAATCGTTTGGACTACTCAGCAGTAAAGACAGATGCAAGT	120
QY	421	GAATAATGTTTACGCAATCTGTATAGCGCAAGTAACGAAAGCAAAAGTAATGATTTCAAGT	480
DB	121	GAATAATGTTTACGCAATCTGTATAGCGCAAGTAACGAAAGCAAAAGTAATGATTTCAAGT	180
QY	481	AGCGTTAGTCTGCACCTATAAAGACAGACACAAAGTGTAGTGTAACTAAACATCGTCA	540
DB	181	AGCGTTAGTCTGCACCTATAAAGACAGACACAAAGTGTAGTGTAACTAAACATCGTCA	240
QY	541	AACACTAATGAGGAAAGAGTGTGGCGCAAAATCCAGACACACAGAAACGACACA	600
DB	241	AACACTAATGAGGAAAGAGTGTGGCGCAAAATCCAGACACACAGAAACGACACA	300
QY	601	TCATCATCAACAAATGCAACTACGGAAGAAACCGCGTAACCTGGTGAAGCTACTACTAG	660
DB	301	TCATCATCAACAAATGCAACTACGGAAGAAACCGCGTAACCTGGTGAAGCTACTACTAG	360

```
QY 661 ACACGATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 720
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 RCAACGATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATGCGGAGAA 420
QY 721 TTAGTGAAATCAAAACAAGTAATGAACGACCTTTTAAATGATACATAATACAGTATCATCTGTA 780
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 TYAGTGAAATCAAAACAAGTAATGAACGACCTTTTAAATGATACATAATACAGTATCATCTGTA 480
QY 781 AATTCACTCAAAATCTTCAAAATGCGGAAATGTTTCAACACGCAAGATCACTTCAACT 840
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 AATTCACTCAAAATCTTCAAAATGCGGAAATGTTTCAACACGCAAGATCACTTCAACT 540
QY 841 GAACACACCTTCAAAACAAGTAATGAGTCCACAGAGTACAGATGCAAGTAATAAGAT 900
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 GAACACACCTTCAAAACAAGTAATGAGTCCACAGAGTACAGATGCAAGTAATAAGAT 600
QY 901 GTAGTTAATCAAGCGGTTTAAATACAGTGGCTGCTAGATGAGAGCAATTTAGTTTACGGCA 960
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 GTAGTTAATCAAGCGGTTTAAATACAGTGGCTGCTAGATGAGAGCAATTTAGTTTACGGCWA 660
QY 961 GTAGCTGAGATGACACCGGAGCTGGACAGATATTACGAATCAGTTCACGAATGTGACA 1020
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 GTAGCTGAGATGACACCGGAGCTGGACAGATATTACGAATCAGTTCACGAATGTGACA 720
QY 1021 GTTGGTATTGACTCTGGTACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAAT 1080
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
721 GTTGGTATTGACTCTGGTACGACTGTGTATCGCACCAAGCAGGTTATGTCAAACTGAAT 780
QY 1081 TATGTTTTCAGTGGCTTAAATCTGCTGTTTAAAGTGACACATTCACAAATCACTTACT 1140
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 TATGKTCTTCACTGTCCTAAATCTGCTGTTTAAAGTGACACATTCACAAATCACTTACT 840
QY 1141 AAGAAATTAATTAATGTTGTTAACTTCACTGCTTAAAGTGACACATTCACAAATCACTTACT 1200
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841 AAGAAATTAATTAATGTTGTTAACTTCACTGCTTAAAGTGACACATTCACAAATCACTTACT 900
QY 1201 GATCAAGTATTGGCAAAATGGTCTAATCGATAGTGTGTAATGTTTATTAATCACTTACT 1260
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
901 GATCAAGTATTGGCAAAATGGTCTAATCGATAGTGTGTAATGTTTATTAATCACTTACT 960
QY 1261 GACTATGTAATTAATTAATGATGTAATAAGCAACTTTGACCATCCCGCTTATATTGAC 1320
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
961 GACTATGTAATTAATTAATGATGTAATAAGCAACTTTGACCATCCCGCTTATATTGAC 1020
QY 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGCGATAGTAGTACACA 1380
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1021 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGCGATAGTAGTACACA 1080
QY 1381 GCACACAAACAGTATTAGTAGATTATGAAATATGTTAAGTATTAATCACTTACTATT 1440
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1081 GCACACAAACAGTATTAGTAGATTATGAAATATGTTAAGTATTAATCACTTACTATT 1140
QY 1441 AAAGGTACAATTAATGACCAATCGATAAACAATAATACGTATCGTCAGACAAATTTATGTC 1500
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1141 AAAGGTACAATTAATGACCAATCGATAAACAATAATACGTATCGTCAGACAAATTTATGTC 1200
QY 1501 AATCCAAAGTGAGATTAAGTATTGCGCGGTTTAAAGGTAATTTAAACCAATACG 1560
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201 AATCCAAAGTGAGATTAAGTATTGCGCGGTTTAAAGGTAATTTAAACCAATACG 1260
QY 1561 GATAGTAATGCAATTAATAGAT - CAGCAAAATCAAGTATTAAAGTATATAAAGTAGATAA 1619
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1261 GATAGTAATGCAATTAATAGATGCAAGCAAAATACAGTATTAAAGTATATAAAGTAGATAA 1320
QY 1620 TGCAGCTGATTTATCGAAGTTACTTTGTGAATCCAGAAATCTTTGAGGATGTCACTAA 1679
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1321 TGCACGAGATTTCTGTAAGTTATGTAATCCAGAAATCTTTGARGATGTCACTRA 1380
QY 1680 TAGTGTGAATTTATCAATTCCTCAATCCAAATCAATTAAGTAGAGTTTAAATACGCTGA 1739
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1381 TAGTGTGAATTTATCAATTCCTCAATCCAAATCAATTAAGTAGAGTTTAAATACGCTGA 1440
QY 1740 TGATCAAAATTACACACCGTATATAGTAGTTGTTAATGTTGATTCATTTGATCCGGAATAGCAA 1799
```

```
Db 1441 TGATCAAAATWACACACCTATATGATGTTTAAATGGKATATTGATCCKAATAGTAA 1500
QY 1800 AGGTGATTAGCTTTAGCTTCAACTTTTATATGGGTATAAATCGAATAATAATTTGCGCTC 1859
Db 1501 AGGTGATTAGCTTTAGCTTCAACTTTTATATGGGTATATGRTATRAYTCRARKWTWTGGMGTC 1560
QY 1860 TATGTCATGGGACACAAAGTAGCATTTAATAACGGATCAGTTCTGTGTGACGGTATCGA 1919
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1561 WATGTCATGGGAAAYAGTAGCATTTAATAACGGATCAGTTCTGTGTGACGGTATCGA 1620
QY 1920 TAAACCAAGTGTCTTCTGAAACAACTGATGAG 1950
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1621 TAAACCTGTGTTCTTCTGAAACAACTGATGAG 1651

RESULT 2
US-60-615-573-18330
; Sequence 18330, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18330
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-60-615-573-18330

Query Match 39.8%; Score 1393.6; DB 8; Length 1600;
Best Local Similarity 91.9%; Pred. No. 2.6e-296;
Matches 1471; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 301 ATGAATATGAAGAAAGAAAGAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTCA 360
Db 1 ATGAATATGAAGAAAGAAAGAAACACGCAATTCGAAAAAATCGATTGGCGTGGCTTCA 60
QY 361 GTGCTTTAGTGTACGTTAATCGTTTGGACTACTCAGCAGTAAAGAGCAGATCAAGT 420
Db 61 GTGCTTTAGTGTACGTTAATCGTTTGGACTACTCAGCAGTAAAGAGCAGATCAAGT 120
QY 421 GAAAAATAGTGTTCGCAATCTGATAGCGCAAGTAAAGAAAGCAAAAGTAATGATTCAGT 480
Db 121 GAAAAATAGTGTTCGCAATCTGATAGCGCAAGTAAAGAAAGCAAAAGTAATGATTCAGT 180
QY 481 AGCGTTAGTGTGCACTTAAACACAGCACAACAGTGTAGTACTAAAAACATCGTCA 540
Db 181 AGCGTTAGTGTGCACTTAAACACAGCACAACAGTGTAGTACTAAAAACATCGTCA 240
QY 541 AACACTTAATAATGCGGAAACAGTGTGCGCAAAATCCAGCACAACAGGAAACGACAAA 600
Db 241 AACACTTAATAATGCGGAAACAGTGTGCGCAAAATCCAGCACAACAGGAAACGACAAA 300
QY 601 TCATCATCAACAAATGCAACTACCGGCAAGAACCGCGTAACTGGTGAAGCTACTACTAG 660
Db 301 TCATCATCAACAAATGCAACTACCGGCAAGAACCGCGTAACTGGTGAAGCTACTACTAG 360
QY 661 ACACGATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGAA 720
Db 361 ACACGATCAAGCTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGGAGGAA 420
QY 721 TTAGTGAATCAAAACAAGTAATGAACGACCTTTTAAATGATACATAATCACTATCTGTA 780
Db 421 TTAGTGAATCAAAACAAGTAATGAACGACCTTTTAAATGATACATAATCACTATCTGTA 480
QY 781 AATTCACTCAAAATCTTCAAAATGCGGAAATGTTTCAACACGCAAGATCACTTCAACT 840
```


Db 481 AATTCACTCAAAATTCACAAATCGGAAAATGTTCAACAACGCAAGATCTCAACT 540
QY
Db 841 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGAGTACAGATCGAAGTAATAAGAT 900
QY
Db 541 GAAGCAACACCTTCAAAACAAATGAATCAGCTCCACAGATACAGATCGAAGTAATAAGAT 600
QY
Db 901 GTAGTTAATCAAGCGGTTAATACAAAGTGCAGCTAGATGAGAGATTTAGTTTACGGCA 960
QY
Db 601 GTAGTTAATCAAGCGGTTAATACAAAGTGCAGCTAGATGAGAGATTTAGTTTACGGCA 660
QY
Db 961 GTAGCTCAGATGACACCGGAGCTGGACAGATATACGAATCAGTTGACGAATGTGACA 1020
QY
Db 661 GTAGCTCAGATGACACCGGAGCTGGACAGATATACGAATCAGTTGACGAATGTGACA 720
QY
Db 1021 GTTGGTATTGACTCTGCTGACACTGTGTATCCGACCAAGCAGGTATGTCAAACTGAAT 1080
QY
Db 721 GTTACATTGACTCTGCTGACACTGTGTATCCGACCAAGCAGGTATGTCAAACTGAAT 780
QY
Db 1081 TATGGTTTTTCAAGTCTAATCTGCTGTTAAAGGTGACACATTCAAAATAACTGTACCT 1140
QY
Db 781 TATGGTTTTTCAAGTCTAATCTGCTGTTAAAGGTGACACATTCAAAATAACTGTACCT 840
QY
Db 1141 AAGAAATTAACCTTAATGTTGTTAACTTCAACTCTAAAGTGCACCAATTAATGGCTGGA 1200
QY
Db 841 AAGAAATTAACCTTAATGTTGTTAACTTCAACTCTAAAGTGCACCAATTAATGGCTGGA 900
QY
Db 1201 GATCAAGTATTGGCAAAATGTTGAATCGATGATGATGATGATGATGATGATGATGATGAT 1260
QY
Db 901 GATCAAGTATTGGCAAAATGTTGAATCGATGATGATGATGATGATGATGATGATGATGAT 960
QY
Db 1261 GACTATGTAATACTAAAGATGATGTTAAAGCAACTTTGACCTGCGCTTATATTGAC 1320
QY
Db 961 GACTATGTTGATAATAAGAAATGTTAAAGCAACTTTGACCTGCGCTTATATTGAC 1020
QY
Db 1321 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATPAGGTAGTACAA 1380
QY
Db 1021 CCTGAAATGTTAAAGACAGGTAATGTGACATTTGGCTACTGGCATPAGGTAGTACAA 1080
QY
Db 1381 GCAAAACAAACAGTATTAGTATGATGTTAAAGATGTTAAAGTATGTTTATTAATCTTCTATT 1440
QY
Db 1081 GCTAGTAAAGCAGTATTATCGACTATGAGAAATATGGAACAATTCCTAATAATTTATCAAT 1140
QY
Db 1441 AAGAGTCAATTCACCAATTCGATAAACAATAATACGATGCTGACAGCAATTTATGTC 1500
QY
Db 1141 AAGAGTCAATTCGATAAACAATAATACGATGCTGACAGCAATTTATGTC 1200
QY
Db 1501 AATCCAGTGGAGATACGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAAAATACG 1560
QY
Db 1201 AATCCAGTGGAGATACGTTATGCGCGGTTTAAACAGGTAAATTTAAACCAAAATACG 1260
QY
Db 1561 GATAGTAAATGATTAATAGATCAGCAAAATACAGTATTAAGTATTAAGTATTAAGTATTAAT 1620
QY
Db 1261 AAGAGTAAATGATTAATAGATCAGCAAAATACAGTATTAAGTATTAAGTATTAAGTATTAAT 1320
QY
Db 1621 GCAGCTGATTTATCTGAAGTTACTTTGTGAATCCAGAAACTTTGAGGATGTCACCTAAT 1680
QY
Db 1321 GCTAATGATTTATCTGAAGTTACTTTGTGAATCCCTAGCGATTTTGAAGATGTAATTAAT 1380
QY
Db 1681 AGTGTGAATTTACATTCACCAATTCGAATCAATATTAAGTATTAAGTATTAAGTATTAAGTATTAAT 1740
QY
Db 1381 CAAGTTAGAAATTTCAATTCGAATGCTTAATCAATCAAAAGTAGAAATTTCTCAGGACGAT 1440
QY
Db 1741 GATCAAAATACACACCGTATATAGTATGTTTAAATGGTCAATTTGATCGAATACGAA 1800
QY
Db 1441 GACCAAAATACACACCGTATATAGTATGTTTAAATGGTCAATTTGATCGTATGATCA 1500
QY
Db 1801 GTGTATTAGCTTTAGCTTCAACTTTATATGGGTATTAACCTCGAATATTAATTTGCGCTCT 1860
QY
Db 1501 GTGTATTAGCTTTAGCTTCAACTTTATATGGGTATTAATTTATGATTTCAATTTATGAGATCT 1560
QY
Db 1861 ATGTCTGGGAACGAAAGTAGCATTTAATAACGGATCAG 1900

Db 1561 ATGTCTGGGAACGAAAGTAGCATTTAATAACGGATCAG 1600

RESULT 3

US-60-615-573-14144/c
; Sequence 14144, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14144
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-60-615-573-14144

Query Match 6.2%; Score 217.8; DB 8; Length 404;

Best Local Similarity 78.5%; Pred. No. 4.2e-39;

Matches 252; Conservative 6; Mismatches 63; Indels 0; Gaps 0;

QY 2357 ATTCGAGTCCAGACAGCGATTCAGATTCGACAGTGTCCGACTCAGATAGCGATTCCG 2416

Db 327 ATTCGAGTATGATTTAGATACAGATATCGTTTCAATAGTACTCAGAAAAATGACACAT 268

QY 2417 ACTCAGATCCGACTCAGATTCAGACAGCGATTCAGATTCAGACAGCGATTCAGATTTCAG 2476

Db 267 ATTTAGATAGTGAATTCAGACTCAGATAGTACTAGATTCAGATAGTGAATTCAGATTTCAG 208

QY 2477 ATAGCGATTCAGATTCGACAGTACTCAGATTCGACAGTGTCCGACTCAGATAGCGATTCAGATAGCG 2536

Db 207 ATAGTATTCAGATTCAGACAGTACTCAGATTCAGATAGTGAATTCAGATTTCAGATTTCAG 148

QY 2537 ATTCAGATTCGACAGTACTCAGATTCGACAGTGTCCGACTCAGATAGCGATTCAGATAGCGATTCAG 2596

Db 147 ACTCAGATTCAGACAGTGTTCAGATTCAGATAGTGAATTCAGATTCAGACAGTGTTCAG 88

QY 2597 ATTCAGCGATGATTCGATTCAGATTCGACAGTGTCCGACTCAGATAGCGATTCAGATTCAGATTCAG 2656

Db 87 ACTCAGATGATTCAGATTCAGATTCAGATAGTGAATTCAGATTCAGATAGTGAATTCAGATTCAG 28

QY 2657 ATAGCGACTCAGACTCGGATA 2677

Db 27 ATAGTATTCAGACTCTGGTA 7

RESULT 4

US-60-615-573-18462/c
; Sequence 18462, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18462
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-60-615-573-18462

Query Match 4.8%; Score 169.4; DB 8; Length 180;

Best Local Similarity 96.6%; Pred. No. 1.13e-28;
Matches 173; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2893 TCAGGTTCTAACAAATAATGTAGTTCGCGCTAAATTCACCTAAAAAATGGTACTAATGCTTCT 2952
Db 179 TCAGGTTCTAACAAATAATGTAGTTCGCGCTAAATTCACCTAAAAAATGGTACTAATGCTTCA 120

QY 2953 AATAAAAAATGAGCCTAAAGTAGTAAGAACCAATTAACAGATACAGGTTCTGAAGATGAA 3012
Db 119 AATAAAAAATGAGCCTAAAGTAAGAAATTAAGAACCAATTAACAGATACAGGTTCTGAAGATGAA 60

QY 3013 GCAAAATACGTCACCTAAATTCGCGGATTATTAGCATCAATAGGTTTCATTACTATTTTCAG 3071
Db 59 GCTAATAACGTCACCTAAATTCGCGGATTATTAGCATCAATAGGTTTCATTACTATTTTCAG 1

RESULT 5
US-10-956-157-5134
; Sequence 5134, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956.157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5134
; LENGTH: 4050
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-5134

Query Match 4.0%; Score 141.4; DB 6; Length 4050;
Best Local Similarity 48.6%; Pred. No. 4.8e-22;
Matches 422; Conservative 0; Mismatches 441; Indels 6; Gaps 1;

QY 2011 AGCGATTCTAATTCAGATAGCGGTTTCAGATTTCGGGTAGTGATTCTACATCAGATAGTGGT 2070
Db 1653 AGGAGTGTGGTTTCATCTAGGAGATGTGGCCCATCTAGGGAATGTGGTTTCATCTAGGAG 1712

QY 2071 TCAGATTACAGCAGTGATTTCAGATTACAGAAAGTGAATTCAGACTCAGCGAGTGATTCAGAT 2130
Db 1713 CGTGGCCCATCTAGGAGCGTGGTTTCATCTAGGAGCGTGGCCCATCTAGGAGCGGTGGC 1772

QY 2131 TCAGAACCGATTCCGACTCAGCGAGCGATTCGCACTCAGACATGACTCGGATTTCAGAT 2190
Db 1773 CCATCTAGGAGCGTGGTTTCATCTAGGAGCGTGGCCCATCTAGGAGATGTGGTTTCATCT 1832

QY 2191 AGCGATTCTGACTCAGACAGTGACTCAGATTCCGACAGTGACTCAGATTTCAGATPAGCAT 2250
Db 1833 AGGAGCGTGGCCCATCTAGGGAATGTGGTTTCATCTAGGAGCGTGGCCCATCTAGGAG 1892

QY 2251 TCTGACTCAGACAGTGACTCAGATTTCAGATTTCAGATPAGCATTCAGATTTCAGATAGCGATTTCAGAT 2310
Db 1893 GGTGGCTCATCTAGGAGCGTGGTTTCATCTAGGAGCGATGGCCCATCTAGGAGCGATGGC 1952

QY 2311 TCCGACAGTGATTCCGACTCAGACAGCGATTTCGACTCCGACAGTGATTTCGCACTCAGAC 2370
Db 1953 CCATCTAGGAGCGTGGCCCATCTAGGAGCGATGGCCCATCTAGGAGCGTGGCTCATCT 2012

QY 2371 AGCGATTTCAGATTCCGACAGTGATTTCGCACTCAGATAGCGATTTCGCACTCAGATAGCGAC 2430
Db 2013 AGGAGCGTGGTTTCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGCCCATCTAGGAG 2072

QY 2431 TCAGATTTCAGACAGCGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGATAGCGATTTCAGAT 2490
Db 2073 CGTGGTTTCATCTAGGAGCGTGGCCCATCTAGGAGCGTGGTTTCATCTAGGAGTGTGGC 2132

QY 2491 TCCGACAGTGACTCAGATTTCGACAGTGACTCGGATTTCAGATPAGCGATTTCAGATTCGAC 2550

```
; CURRENT APPLICATION NUMBER: US/10/746,294A
; CURRENT FILING DATE: 2003-12-29
; PRIOR APPLICATION NUMBER: US 09/692,412
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 09/803,736
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 124
; SEQ ID NO 30
; LENGTH: 98412
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-746-294A-30

Query Match      2.4%; Score 85; DB 6; Length 98412;
Best Local Similarity 48.0%; Pred. No. 3.2e-09;
Matches 275; Conservative 0; Mismatches 295; Indels 3; Gaps 1;

QY 2115 AGCGAGTGAATTCAGATTTCAGAAAGCGGATTCGACTCAGCGAGCGATTCCGACTCAGACAA 2174
Db 51443 AGAGAGAGAACTAGAGGCTGCAGGTGGCTGAGAGGGACAGGTTGATTGAAATGTCGGTGG 51384

QY 2175 TGACTCGGATTTCAGATAGCGAATTCGACTCAGACAGTGAATTCAGATTCCGACAGTGAATC 2234
Db 51383 TGGCTGAGACACTGGCAACCGCTTAGAGGTTGTAGGTGGCTGAGATACCGGCAATGGCTT 51324

QY 2235 AGATTTCAGATAGCGATTCTGACTCAGACAGTGAATTCAGATTTCAGATAGCGAATTCAGATTC 2294
Db 51323 AGAGGATTTCGGGTGGCTGAGACACA---AATGGCTTAGAGGTCGTTGGTGGCTGAGAGAT 51267

QY 2295 AGATAGCGATTTCAGATTCCGACAGTGAATTCGACTCAGACAGCGATTCTGACTCCGACAG 2354
Db 51266 AGGGAATGACTTAGAGGCGCTGGTGGCTGAGATGCTGGAAAGGCTTAGCGTGCATGA 51207

QY 2355 TGATTCCGACTCAGACAGCGAATTCAGATTCCGACAGTGAATTCGACTCAGATAGCGAATTC 2414
Db 51206 TGGTTGAGACATCGAATAATGGCTGAGAAATTCGTTGGTGGGACACCGGGAACGGGTT 51147

QY 2415 CGACTCAGATAGCGATTCAGATTTCAGACAGCGAATTCAGATTTCAGACAGCGAATTCAGATTC 2474
Db 51146 AGAAGCGGGTTGGCTGAGACACCGGAAATGGCTTAGAGTTTGGATGGCTGAGACAC 51087

QY 2475 AGATAGCGATTTCAGATTCCGACAGTGAATTCGACAGTGAATTCGAGTTCGATTCAGATAG 2534
Db 51086 CGGGAATGGCTTAGAGTTTGTATGACTGAGAGACCGGCAATGGCTTTGACCTTTGGA 51027

QY 2535 CGATTTCAGATTCCGACAGTGAATTCAGATTCCGACAGTGAATTCGACTCAGACAGTGAATTC 2594
Db 51026 TGACTGAGAAACCGACAAAGGCTTAGACITTTTGTGATGACTGAGACACTGGGAATGGCTT 50967

QY 2595 GGATTTCAGCGAGTGAATTCGGAATTCAGATAGTGAATTCGACTCCGACAGTGAATTCGAGTTC 2654
Db 50966 AGATTTCGATGACTGAGACATCGGAACGGCTTAGAGTTTCGATGGTGGCTGAGACAA 50907

QY 2655 AGATAGCGACTCAGACTCGGATAGCGACTCGGA 2687
Db 50906 TGGAAATGGTTAGAGGTCGGTGGCTGAGA 50874

RESULT 8
US-60-615-573-18342
; Sequence 18342, Application US/60615573
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; APPLICANT: Murphy, Ellen
; APPLICANT: Olmsted, Stephen
; TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
; FILE REFERENCE: 031896-084099 (AM 101724)
; CURRENT APPLICATION NUMBER: US/60/615,573
; CURRENT FILING DATE: 2004-10-05
; NUMBER OF SEQ ID NOS: 18598
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18342
; LENGTH: 2701
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-60-615-573-18342
```

```
Query Match      2.3%; Score 79.8; DB 8; Length 2701;
Best Local Similarity 43.3%; Pred. No. 1.1e-08;
Matches 393; Conservative 25; Mismatches 477; Indels 12; Gaps 3;

QY 606 ATCAAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 665
Db 468 AGCACCATAATWATACAAATTAACGATGTAATCAAAACCATCTCAAGTGAACCATCTAC 527

QY 666 GAATCAAGCTTAATACACCGGCAACAACTCAATCAAGCAATACAAATCGCGAGGAATTAGT 725
Db 528 AAGTGAATTCAAACCAAAACCAACTACCTCAAGAAATCTACAAATATTGAAATTCACA 587

QY 726 GAATCAAAACAAGTAATGAACGACTTTTATGATGATCTAATACAGATATCATCTGTAATTC 785
Db 588 ACCGCAACCAACCGCTTCAAAAGTAGACAATCAAGTTACAGATGCAACTAATCCAAAGA 647

QY 786 ACCTCAAAATTTCAAAATGCGGAAATGTTTCAACAAACGCAAGATACCTTCAACTGAAGC 845
Db 648 ACCAGTAATGTTCTCAAAAGAGAACTTAAANAAATCTTGAGAAATTAAGAAATTTGT 707

QY 846 AACACCTTCAACAAATGAATCAGCTCCACAGATGACAGATGCAAGTAATTAAGATGATG 905
Db 708 TAGAAATGATFARCAATACAGATCTTTCACTAAACCAAGTTGCTTACAGCTCCAACAGT 767

QY 906 TAATCAAGCGGTTAATACAGTGGCTAGAAATGAGAGCAATTTAGTTTAGCGGCAAGTAGC 965
Db 768 TGCACCAAAACGKTAATGVCRAAATGCGYTTTGCAAGTTGCAACACGAGCAGCTGTC 827

QY 966 TGCAGATGCAACCGGCAAGCTGGCAAGATTTAGCAATCAGTTGCAAGTAATGACAGTTGG 1025
Db 828 TTCAAAATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 887

QY 1026 TATGACTCTGGTACGACTGTGTATCCGACCAAGCAGGTTATGTCAAACTGAATTAATGG 1085
Db 888 TAAAGATAATGTGGCAGCAG-----CGCATGACGGTAAAGATATTGAATATGATACAGA 941

QY 1086 TTTTTCAGTGCCTTAATTCGCTGTTAAAGGTGACACATTCAAATTAATTAATTAATTAATTA 1145
Db 942 GTTTCAATTTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1001

QY 1146 ATTAACCTTTAAATGGTGTAAAC---TTCAACTGCTAAAGTGCACCAATTTATGGCTGGAGA 1202
Db 1002 TGTAAATTCCTCGGATTTAAACAGATAAAATGATCCTATCGATATTACTGATCCATCAGG 1061

QY 1203 TCAGTATTGGCAATATGGTGTATTCGATAGTGTATGTAAGTAAAGTAAAGTAAAGTAAAGT 1259
Db 1062 AGAGGTTCATTCGTAAGGAACATTTGATTAAGGCACTAAGCAAAATCACATATACATTTAC 1121

QY 1260 AGACTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1319
Db 1122 AGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1181

QY 1320 CCTGAAATTTTAAAGAGACAGGTAATGTCACATTTGGCTACTGGCATAGGTAGTACAAAC 1379
Db 1182 TAARMAARCAGTWCWAAATGACACAGGTTTGAATTTAATTTTTCGACAGCAGGTAAGA 1241

QY 1380 AGCAAAACAAACAGTATTAGTATGATTAAGAAAAATATGTAAGTATTTTAACTTATCTAT 1439
Db 1242 AACWAGCCAAAAAGTTCGTTGATTTATCAAGAYCCAATGGTYCATGTTGATTTCAACAT 1301

QY 1440 TAAAGGTACAATTTGACCAATCGATAAACAATTAATTAAGTATCGTACAGCAATTTATGT 1499
Db 1302 TCAATCTATCTTTACAAATTTAGATGAARAYAAACAACTATTGAACACAAATTTATGT 1361

QY 1500 CAATCCA 1506
||:|:
```


1421 AGTTTATACTTATCTATTAAAGGTACAAATTGACCAAAATCGATAAAACAAATAATACGT 1480
1292 AAAATCGGACTGTAACCTTCAAGATATGTTTACAAATATAGATACGAAATTAATACAG 1351
1481 ATCGTCAGACAAATTTATGTCATCAATCAAGTGGAGATAACGTTATTGCGCCGGTTTAAACAG 1540
1352 TTGAGCAAAAGGATTTATATTAATTAACCTCTCTGTTATTCAGCAAGAAACAAATGTA---- 1407
1541 GTAATTTAAACCAAAATACGATAGTAAATGCAATTAATAGATACGCAAAATACAAGTATTA 1600
1408 --AATATTTTTCAGGAATGGGATGAAGGTTCAACAATTTATGACGATAGACAAATTAATTA 1465
1601 AAGTATATAAAGTAGATAAATGCAGCTGATTTATCTGAAAGTTACTTTGTGAATCCAGAAA 1660
1466 AAGTTTATAAGTTGGAGATAATCAAAATTTTACCAGATAGTAACAGAAATTTATGATTAACA 1525
1661 ---ACTTTGAGGATGTCACHTAATAGTGTGAATATTACATTTCCCAATCCAAATCAATATA 1717
1526 GTCAATATGAAGATGTCACAAATGATGATTATGCCCCAATTAGGAAATTAATGAYGTGA 1585
1718 AAGTAGATTTTAATACCCCTGATGATCAAAATTA 1750
1586 ATATTAAATTTTGGTAATATAGATTCACCAATATA 1618

RESULT 12

US-60-613-292-360

; Sequence 360, Application US/60613292

; GENERAL INFORMATION:

; APPLICANT: Gene Logic, Inc

; APPLICANT: HIGGS, Brandon

; APPLICANT: ELASHOFF, Michael

; APPLICANT: MENDRICK, Donna L.

; APPLICANT: PORTER, Mark W.

; APPLICANT: CASTLE, Arthur L.

; APPLICANT: JOHNSON, Kory R.

; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES

; FILE REFERENCE: 044921-5135-PR

; CURRENT APPLICATION NUMBER: US/60/613,292

; CURRENT FILING DATE: 2004-09-28

; NUMBER OF SEQ ID NOS: 1300

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 360

; LENGTH: 5192

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; OTHER INFORMATION:

US-60-613-292-360

Query Match 2.0%; Score 71.6; DB 8; Length 5192;

Best Local Similarity 52.7%; Pred. No. 1e-06;

Matches 155; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

2429 ACTCAGATTCAGACAGCGATTCAGATTCACACAGCGATTCAGATTCAGATTCAGATTCAG 2488
3785 ACACAGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGATAGAC 3844
2489 ATTCGACAGTACTCAGATTCGACAGTACTCGGATTCAGATTCAGATTCAGATTCG 2548
3845 AGATAGATGACAGAGAGAGCGGACAGAGACAGACAGAGAGAGAGAGACAGAGAAAGAG 3904
2549 ACAGTACTCAGATTCGACAGTACTCAGATTCAGACAGTACTGATTCGATTCAGCGAGTG 2608
3905 AGAGAGAGAGAGAGAGAGACAGACAGACAGACAGACAGACAGACAGAGAGAGAGAGAGAG 3964
2609 ATTCGATTCAGATAGTATTCGATTCGACAGTACTCGGATTCAGATTCAGATTCAGACTCAG 2668
3965 ATAGACAGACAGACAGACAG 4024
2669 ACTCGGATAGCGACTCGGATTCAGATTCAGATTCGATTCGATTCGATTCGATTCGATTCAG 2722

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2004, 05:31:43 ; Search time 10221 Seconds
(without alignments)
12471.009 Million cell updates/sec

Title: US-09-679-643-1

Perfect score: 3498

Sequence: 1 ggtaaccataataacacatc.....gaagagtataagaagaagctt 3498

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	740.8	21.2	768	9	CC958014 BOIHQ26TR
C 2	518.2	14.8	534	9	CC942933 BOIHQ26TF
C 3	212.4	6.1	697	5	BW250931 BW250931
C 4	205.2	5.9	772	8	CC099392 CSU-K34.1
C 5	205	5.9	773	8	CC099378 CSU-K34.1
C 6	182.4	5.2	811	7	CO426942 UI-M-HUO-
C 7	179.8	5.1	667	5	BW254640 BW254640
C 8	171.8	4.9	699	5	BX729701 BX729701
C 9	167.6	4.8	928	9	CL129761 ISB1-97D4
C 10	167.4	4.8	563	7	CN689123 E0270H03-
C 11	164.6	4.7	817	9	CC507028 CH240_348
C 12	163	4.7	625	1	AV992680 AV992680
C 13	162.8	4.7	544	7	CN688762 E0264F04-
C 14	162.2	4.6	973	7	CK423535 AUF-IPSPn
C 15	160.2	4.6	1526	1	AL162088 DRFP7611
C 16	158.6	4.5	607	7	CF894873 A0141B10-
C 17	158.4	4.5	737	8	BZ191186 CH230-424
C 18	157.4	4.5	870	5	BX078600 BX078600
C 19	156.8	4.5	542	8	AZ875941 2M0190D14
C 20	156.6	4.5	623	8	CC075747 CSU-K33r.
C 21	155.2	4.4	542	8	AZ875941 2M0190D14
C 22	152.4	4.4	941	7	CK425697 AUF-IPtes
C 23	150.6	4.3	599	6	CB582745 AMGNNUC:N
C 24	149	4.3	652	5	BW388427 BW388427

25	148.8	4.3	944	9	CNS04SSE	AL305591 Tetraodon
C 26	148.6	4.2	573	7	CF893788	A0127F06-
C 27	146.8	4.2	744	5	EX078599	EX078599
C 28	146.6	4.2	860	9	CNS02E27	AL193192 Tetraodon
C 29	146	4.2	564	1	AV854983	AV854983
C 30	145.2	4.2	927	7	CK420227	CK420227 AUF-IPtrk
C 31	145	4.1	704	7	CF899579	CF899579 A0304H06-
C 32	141.2	4.0	680	7	CK344793	CK344793 K0975C05-
C 33	141.2	4.0	984	9	CNS071N7	AL425141 clone BAO
C 34	140.6	4.0	650	9	AG273016	AG273016 Cyanidios
C 35	140.4	4.0	623	8	CC075747	CC075747 CSU-K33r.
C 36	140	4.0	950	7	CK425370	CK425370 AUF-IPtes
C 37	139.4	4.0	832	4	BM170146	BM170146 EST572669
C 38	139.2	4.0	723	9	CNS01UN74	AL168241 Tetraodon
C 39	137.8	3.9	1939	3	AK079258	AK079258 Mus muscu
C 40	136.8	3.9	571	8	BZ214163	BZ214163 CH230-331
C 41	136.2	3.9	773	6	CB939384	CB939384 IPCGX14
C 42	136	3.9	764	2	BF024332	BF024332 Pap 058 L
C 43	135.4	3.9	960	9	CNS01EB6	AL140431 Anopheles
C 44	135.2	3.9	891	5	BM120795	BM120795
C 45	134.4	3.8	554	6	CB613746	CB613746 AMGNNUC:N

ALIGNMENTS

RESULT 1
CC958014/c
LOCUS
CC958014 BOIHQ26TR BO 1.4 1.6 KB nuc Brassica oleracea genomic clone
DEFINITION
BOIHQ26, genomic survey sequence.
ACCESSION
CC958014 GI:33801933
VERSION
GSS.
KEYWORDS
Brassica oleracea
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
REFERENCE
Town, C.D., Van Aken, S., Usterback, T., Koc, H. and Fraser, C.M.
AUTHORS
Whole genome shotgun sequencing of Brassica oleracea
TITLE
Unpublished (2001)
JOURNAL
Other GSSs: BOIHQ26TF
COMMENT
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1 . . 768
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BO 1.4 1.6 KB nuc"
/note="Vector: pHOS2; Site 1: BstXI; 1.4-1.6 kb sheared
nuclear DNA inserted into pHOS2 using BstXI linkers"

FEATURES

Query Match 21.2%; Score 740.8; DB 9; Length 768;
Best Local Similarity 97.8%; Pred. No. 1.1e-138;
Matches 751; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 1275 TAAAGATGATGTAAGACCTTTTGACCATGCGCGCTTATATTGACCTGAAATCTTAA 1334
Db 768 TAAAGATGATGTAAGACCTTTTGACCATGCGCGCTTATATTGACCTGAAATCTTAC 709

1335 AAAGACAGTAAATGTGACATTTGGTCTACTGCGCATAGTGTAGTCAACAGCAAAACAAAACAGT 1394
 1395 ATTAGTAGATATGAAAAATATGTAAGTTTATTAACCTTATCTATTAAGGTGACAAATGGA 1454
 648 ATTAGTAGATATGAAAAATATGTAAGTTTATTAACCTTATCTATTAAGGTGACAAATGGA 589
 1455 CCAATCGATTAACAAATATATGATGTCGTCAGCAAAATTTATGTCATCAAGTGGAGA 1514
 588 CCAATCGATTAACAAATATATGATGTCGTCAGCAAAATTTATGTCATCAAGTGGAGA 529
 1515 TAACGTTATGCGCGGTTTAAACAGGTAATTTAAACCAATACCGATAGTAAGTGCATT 1574
 528 TAACGTTATGCGCGGTTTAAACAGGTAATTTAAACCAATACCGATAGTAAGTGCATT 469
 1575 AATAGTAGCAAAATATCAAGTATTAAGTATATTAAGTAGATATTAAGTGTGATTTATC 1634
 468 AATAGTAGCAAAATATCAAGTATTAAGTATATTAAGTAGATATTAAGTGTGATTTATC 409
 1635 TGAAGTTACTTTGTAATCCAGAACTTTGAGGATGTCATTAAGTGTGATTTATC 1694
 408 TGAAGTTACTTTGTAATCCAGAACTTTGAGGATGTCATTAAGTGTGATTTATC 349
 1695 ATTCCCAATCCAAATCAATATTAAGTAGAGTTTAAATACGCTGATGATCAAAATTAACAAC 1754
 348 ATTCCCAATCCAAATCAATATTAAGTAGAGTTTAAATACGCTGATGATCAAAATTAACAAC 289
 1755 ACCGTATATAGTAGTTTAAATGTCATTAAGTGTGATGATGATGATGATGATGATGATGAT 1814
 288 ACCGTATATAGTAGTTTAAATGTCATTAAGTGTGATGATGATGATGATGATGATGATGAT 229
 1815 ACCTTCACTTTATATGGTATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1874
 228 ACCTTCACTTTATATGGTATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 169
 1875 CGAAGTAGCATTTAATAACCGATCAGGTTCTGGTACGATGATGATGATGATGATGATGATGAT 1934
 168 CGAAGTAGCATTTAATAACCGATCAGGTTCTGGTACGATGATGATGATGATGATGATGATGAT 109
 1935 TGAACAACTGATGAGCTGCTGTAATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 1994
 108 TGAACAACTGATGAGCTGCTGTAATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 49
 1995 AGGTCAGATTTGCGCAGGTTCTTAATTCAGATGAGGTTTCAAGATTC 2042
 48 AGGTCAGATTTGCGCAGGTTCTTAATTCAGATGAGGTTTCAAGATTC 1

CC942933 534 bp DNA linear GSS 18-AUG-2003
 LOCUS BOHQ26TF BO_1.4_1.6 KB nuc Brassica oleracea genomic clone
 DEFINITION BOHQ26, genomic survey sequence.
 CC942933
 ACCESSION CC942933.1 GI:33775799
 VERSION GSS.
 KEYWORDS Brassica oleracea
 SOURCE Brassica oleracea
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Eukaryota; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 534)
 Town, C.D., V. Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished (2001)
 Other GSSs: BOHQ26TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: Sheared ends.
 Location/Qualifiers
 1. 534
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOHQ26"
 /clone_lib="BO 1.4_1.6 KB nuc"
 /note="Vector: pHOS2; Site 1: BstXI; 1.4-1.6 kb sheared
 nuclear DNA inserted into pHOS2 using BstXI linkers"

ORIGIN

Query Match 14.8%; Score 518.2; DB 9; Length 534;
 Best Local Similarity 99.3%; Pred. No. 7.4e-94;
 Matches 531; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 418 AGTGAATAATAGTGTACGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTCA 477
 DB 1 AGTGAATAATAGT-TTCGCAATCTGATAGCGCAAGTAAACGAAAGCAAAAGTAATGATTCA 59
 QY 478 AGTAGCGTTAGTGTGACACCTTAAACAGACACAAACGTCGAGTGATCTAAACATCG 537
 DB 60 AGTAGCGTTAGTGTGACACCTTAAACAGACACAAACGTCGAGTGATCTAAACATCG 119
 QY 538 TCAACACTAATAATGCGGAACGAGTGTGCGCAAAATCCAGCACACACGAAACGACA 597
 DB 120 TCAACACTAATAATGCGGAACGAGTGTGCGCAAAATCCAGCACACACGAAACGACA 179
 QY 598 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAATCTGGTGAAGTACTACT 657
 DB 180 CAATCATCATCAACAAATGCAACTACGGAAGAAACGCGGTAATCTGGTGAAGTACTACT 239
 QY 658 ACGCAACGAAATCAAGTAAATACACCGCAACAACTCAATCAAGCAATACAAATCGGAG 717
 DB 240 ACGCAACGAAATCAAGTAAATACACCGCAACAACTCAATCAAGCAATACAAATCGGAG 299
 QY 718 GAATAGTAGAATCAACAAAGTAATGAACGACTTTTATGATCTACTTAATACAGTATCACT 777
 DB 300 GAATAGTAGAATCAACAAAGTAATGAACGACTTTTATGATCTACTTAATACAGTATCACT 359
 QY 778 GTAAATTCACCTCAAAATTTCTCAAAATCGGAAATGTTTCAACAAACGCAAGATCTTCA 837
 DB 360 GTAAATTCACCTCAAAATTTCTCAAAATCGGAAATGTTTCAACAAACGCAAGATCTTCA 419
 QY 838 ACTGAAGCAACACCTTCAAAATCAATGAATCAAGTCCACAGAGTACAGATGCAAGTAAATAA 897
 DB 420 ACTGAAGCAACACCTTCAAAATCAATGAATCAAGTCCACAGAGTACAGATGCAAGTAAATAA 479
 QY 898 GATGTAGTAAATCAAGCGGTTAATACAGTGGCTTAGAATGAGAGCAATTTAGTT 952
 DB 480 GATGTAGTAAATCAAGCGGTTAATACAGTGGCTTAGAATGAGAGCAATTTAGTT 534

RESULT 3
 LOCUS BW250931/c
 DEFINITION BW250931 Nori Satoh unpublished cDNA library, tailbud embryo Clona
 intestinalis cDNA clone citb08h22 5', mRNA sequence.
 CC942933
 ACCESSION BW250931.1 GI:24830849
 VERSION EST.
 KEYWORDS Ciona intestinalis
 SOURCE Ciona intestinalis
 ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.
 1 (bases 1 to 697)
 Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished (2002)
 Contact: Nori Satoh


```
Db 348 GATTGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGG 289
QY 2656 GATACGACTCAGACTCGGATAGCGACTCGGATTCAGATAGCGATTCCGACTCAGATAGC 2715
Db 288 GATTGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGGGATTGG 229
QY 2716 GATTGAGAAATCAGA 2729
Db 228 GATTGGGATTGGGA 215

RESULT 5
CC099378/c
LOCUS
DEFINITION
CSU-K34.113M15.SP6 CSU-K34 Aedes aegypti genomic clone
CSU-K34.113M15, genomic survey sequence.
ACCESSION
CC099378
VERSION
CC099378.1 GI:29965941
KEYWORDS
GSS.
SOURCE
Aedes aegypti (yellow fever mosquito)
ORGANISM
Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;
Stegomyia.
REFERENCE
1 (bases 1 to 773)
Loftus,B., Shetty,J., Severson,D., Brown,S. and Knudson,D.
End sequencing of Aedes aegypti BACs
Unpublished (2003)
JOURNAL
Other GSSs: CSU-K34.113M15.T7
COMMENT
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: entaetigr.org
Library was provided by Susan Brown and Dennis Knudson at Colorado
State University.
Seq primer: SP6
Class: BAC ends.

FEATURES
source
1..773
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/db_xref="taxon:7159"
/clone="CSU-K34.113M15"
/clone_lib="CSU-K34"
/note="Vector: pBACe3.6; Site 1: EcoRI; Source DNA: Aedes
aegypti; strain unknown (derived from freshly hatched
larvae at the Virus Research Centre, Poona, India.
Reference: SINGH, K. R. P., 1967 Cell cultures derived
from larvae of Aedes albopictus (Skuse) and Aedes aegypti
(L.). Current Science 36: 506-508; ATC-10 cell line ATCC
CCL-125"

ORIGIN
Query Match 5.9%; Score 205; DB 8; Length 773;
Best Local Similarity 58.0%; Pred No. 9.2e-31;
Matches 381; Conservative 0; Mismatches 275; Indels 1; Gaps 1;

QY 2266 GACTCAGATTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2325
Db 767 GATTTCAGATTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 708
QY 2326 GACTCAGACAGGATTCGATTCGACAGTGTTCGACTCAGACAGGATTCAGATTC 2385
Db 707 GATTTCAGATTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 648
QY 2386 GACAGTGTTCGACTCAGATTCAGATTCAGATTCGACTCAGATTCAGATTCAGACAGC 2445
Db 647 GATTTCAGATTTCAGATTCAGATTCAGATTCATTTAGATTCAGATTCAGATTCATTA 588
QY 2446 GATTTCAGATTTCAGACAGGATTCAGATTCAGATTCAGATTCAGATTCGACAGTCA 2505
```

```
Db 587 GATTCAAATTTAGATTCAGATTCAGATTCATATTTGAATTCAGATTTAGATTTCTGATTAA 528
QY 2506 GATTCGACACTGACTCGGATTCAGATTCAGATTCAGATTCGACAGTCCGACAGTCC 2565
Db 527 GATTCATATTTAGATTCAGATTCAGATTCAGATTCGATTCGATTCAGATTCAGATTTA 468
QY 2566 GACAGTGTTCAGATTCAGACAGTTCGATTCGATTCGAGTTCGAGTTCGATTCGATTCAGATAGT 2625
Db 467 GACCCAGATTTAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 408
QY 2626 GATTCGAGTCCGACAGTGTTCGATTCAGATTCAGATTCAGATTCGATTCGATTCGATTCG 2685
Db 407 AATTCAAATTTAAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTTA 349
QY 2686 GATTCAGATTCGATTCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2745
Db 348 GATTCAAATTTAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTTA 289
QY 2746 GACAGCGATTCAGATTCAGACAGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCG 2805
Db 288 GACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 229
QY 2806 GATTCAGCGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2865
Db 228 GATTTAGGTTCAAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 169
QY 2866 GAAAGTGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2922
Db 168 CAAACATTTGAAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 112

RESULT 6
CO426942/c
LOCUS
DEFINITION
IMAGE:30674147 5', mRNA sequence.
ACCESSION
CO426942
VERSION
CO426942.1 GI:49673236
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 811)
NIH-MGC http://mgc.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 18-114, >(GAA)nSimple_repeat
Seq primer: pYX-5.
FEATURES
source
1..811
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30674147"
/tissue_type="whole eye"
/dev_stage="newborn ( 1, 5, 15 days )"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HU0"
/note="Organ: Eye; Vector: pYX-Asc; Site_1: EcoR I;
```


LOCUS BX729701 699 bp mRNA linear EST 18-NOV-2003
DEFINITION BX729701 XGC-tadpole Xenopus tropicalis cDNA clone Ttpa075e10 5',
RNA sequence.
ACCESSION BX729701
VERSION BX729701.1 GI:38402442
KEYWORDS EST.
SOURCE xenopus tropicalis (western clawed frog)
ORGANISM xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 699)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: Ttpa075e10.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..699
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="Ttpa075e10"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end"
ORIGIN
Query Match 4.9%; Score 171.8; DB 5; Length 699;
Best Local Similarity 53.3%; Pred No. 4.5e-24;
Matches 362; Conservative 0; Mismatches 317; Indels 0; Gaps 0;
QY 1976 AGGATTCAGATTCGACCGAGTTTCAGATTCGCGAGGATTCCTAAATTCAGATAGCGGTT 2035
DB 20 AGGAATCCATAAAGTCACATGATGCTATTGGGGTCTCTGATGTTGGTACTGGTACTGGTA 79
QY 2036 CAGATTCGGGTAGTGATTCATCATCAGATAGTGGTTCAGATTCACGAGTGAATTCAGATT 2095
DB 80 CTGATACGATACGATGATCGGTACTGGTACTGGTACTGGTACTGGTACTGATCGGTT 139
QY 2096 CAGCAAGTGATTCAGACTCAGCGAGTGATTCAGATTCAGCAAGCGAATCCGACTCAGCGA 2155
DB 140 CTGGTTCGTGCTAGTACTGATGATCTGGTACTGGTACTGATGATCTGGTACTGGTA 199
QY 2156 GCGATTCGGACTCAGACAAATGACTCGGATTCAGATPAGCGATTCGAGTCAAGAGTGAAT 2215
DB 200 CTGATACGTTCTCGTACTGCTGATGATCTGGTACTGGTACTGATGATCTGGTACTGGTA 259
QY 2216 CAGATTCGAGTCACTCAGATTCAGATTCAGATTCGATTCGATTCAGACAGTCACTCAGATT 2275
DB 260 CTGGTACTGATCGGTTCTGGTACTGGTACTGATCGGTTCTGGTACTGATCGGTA 319
QY 2276 CAGATPAGCGATTCAGATTCAGATTCAGATTCGAGAGTGAATTCGAGTCCGACTCAGACA 2335
DB 320 CTGGTACTGATCGGTTCTGGTACTGGTCTGGTACTGATGATGATGATGATGATGATGATG 379
QY 2336 GCGATTCGATCCGACAGTGAATTCGAGTCCGACTCAGACAGCGATTCAGATTCGAGAGTGAAT 2395

DB 380 CTGGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 439
QY 2396 CCGACTCAGATAGCGATTCGAGTCCGACTCAGATCCGACTCAGATTCAGACAGCGATTCAGATT 2455
DB 440 CTGATACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGT 499
QY 2456 CAGACAGCGATTCAGATTCAGATTCAGATTCAGATTCGACAGTCACTCAGATTCGAGACA 2515
DB 500 CTGGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
QY 2516 GTGACTCGGATTCAGATAGCGATTCAGATTCGACAGTGAATTCAGATTCGACAGTGAATTC 2575
DB 560 CTGATACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTA 619
QY 2576 CAGACTCAGACAGTCACTCGGATTCAGCGAGTGAATTCGGATTCAGATGATGATTCGACT 2635
DB 620 CTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTACTGGTA 679
QY 2636 CCGACAGTCACTCGGATTC 2654
DB 680 CTGGTACTGGTACTTATAC 698
RESULT 9
LOCUS CL129761/c 928 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-97D4 T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-97D4,
genomic survey sequence.
ACCESSION CL129761
VERSION CL129761.1 GI:40623396
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 928)
Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
AUTHORS A physical map of the xenopus tropicalis genome
TITLE Unpublished (2003)
JOURNAL Contact: Richard K Wilson
COMMENT Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 11
High quality sequence stop: 723.
FEATURES
Location/Qualifiers
1..928
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-97D4"
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 4.8%; Score 167.6; DB 9; Length 928;
Best Local Similarity 52.9%; Pred. No. 3.2e-23;
Matches 387; Conservative 0; Mismatches 339; Indels 6; Gaps 1;
QY 2135 CAACGGATTCGATTCAGCGAGCGATTCGCGACTCAGACATGACTCGGATTCAGATTCGCG 2194
DB 750 CAATCATACCACTGAGATACCTTACTGAGTACTGAGTACTGAGTACTGAGTACTGAGTACT 691
QY 2195 ATTCTGACTCAGACAGTGAATTCAGATTCGCGAGTGAATTCAGATTCAGATTCAGATTCG 2254
DB 690 ACTGAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 631

2255 ACTCAGACAGTACTCAGATTTCAGATAGCAGATTTCAGATTCAGATTCAGATTCGCG 2314
Db ACTGAGCTACCTACTCAGCTACCTACTGAGCTACCTACTGAGTACCTACTGAGTACCT 571
QY ACAGTGATTCGCGACTCAGACAGCGATTCTGACTCCGACAGTGTATTCGCGACTCAGACAGCG 2374
Db ACTGAGATACCTACTCAGCTACCTACTGAGCTACCTACTGAGTACCTACTGAGTACCT 511
QY ATTGAGATTCGCGAGTGTATTCGAGTTCAGATAGCGATTTCGCGACTCAGATAGCGATTCAG 2434
Db ACTGAGATACCTACTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTAG 457
QY ATTGAGATTCGCGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 2494
Db CTACCTACTGAGATACCTACTGAGATACCTACTGAGTGTATTCAGTGTATTCAGTGTAG 397
QY ACAGTGACTCAGATTCGCGAGTGTATTCAGATAGCGATTTCAGATTCGCGAGTGT 2554
Db ATACCTACTGAGAACTTACGGAGATAAATCTGAGTGTATTCAGTGTATTCAGTGTAG 337
QY ACTGAGATTCGCGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 2614
Db TTACCTACTGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 277
QY ATTGAGATTCGCGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 2674
Db ATACCTACTGAGATACCTACTGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 217
QY ATAGCGACTCGGATTCAGATAGCGATTTCGCGACTCAGATAGCGATTTCAGATTCAGACAGCG 2734
Db ATACCTACTGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 157
QY ATTGAGATTCGCGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 2794
Db ATAACTACTGAGATACCTACTGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 97
QY ATAGTGTACTCGGATTCAGCGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAGTGT 2854
Db ATACCTACTGAGATACCTACTGAGTGTATTCAGTGTATTCAGTGTATTCAGTGTATTCAG 37
QY 2855 ATTCGCGACTCAG 2866
Db ATACCTACTGAG 25

RESULT 10
CN689123/c
LOCUS
DEFINITION
E0270H03-5 NIA Mouse Embryonic Stem (ES) cell (Lif-, 48 h, high density) cDNA library (Long) Mus musculus cDNA clone NIA:E0270H03 IMAGE:30856118 5', mRNA sequence.

ACCESSION
CN689123
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE
AUTHORS
Sharov, A.A., Piao, Y., Matoba, R., Dudekula, D.B., Qian, Y., VanBuren, V., Falco, G., Martin, P.R., Stagg, C.A., Bassey, U.C., Wang, Y., Carter, M.G., Hamatani, T., Aiba, K., Akutsu, H., Sharova, L., Tanaka, T.S., Kimber, W.L., Yoshikawa, T., Jaradat, S.A., Fanfani, S., Nagaraja, R., Boheler, K.R., Taub, D., Hodes, R.J., Longo, D.L., Schlessinger, D., Keller, J., Klotz, E., Kelsae, G., Umezawa, A., Vescovi, A.L., Rossant, J., Kunath, T., Hogan, B.L., Curci, A., D'Urso, M., Kelso, J., Hide, W. and Ko, M.S.

TITLE
Transcriptome analysis of mouse stem cells and early embryos
JOURNAL
PLoS Biol. 1 (3), 410-419 (2003)
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: E0270 row: H column: 03
Seq primer: M13 Reverse
High quality sequence stop: 563
POLYA=No.

FEATURES
Source

1. 563
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129Sv/EvTac"
/db_xref="niaEST:E0270H03-5"
/db_xref="taxon:10090"
/clone="NIA:E0270H03 IMAGE:30856118"
/sex="Male"
/tissue_type="Embryonic Stem Cell"
/cell_line="129.3 ES cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Embryonic Stem (ES) cell (Lif-, 48 h, high density) cDNA library (Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library [Ref. Genome Res. 11: 1553-1558 (2001). PMID: 11544199]. ES cells were plated at density 3x104/cm2 on gelatin-coated plates and cultured for 48 hrs at 37 OC, 5% CO2. Culture medium: DMEM supplemented with 15% FBS, 2 mM L-glutamine, 0.1 mM NEAA, 1mM Sodium pyruvate, 0.1 mM beta-mercaptoethanol, 100 U/ml penicillin, and 100 ug/ml streptomycin. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]. 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3' from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.4 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 4.8%; Score 167.4; DB 7; Length 563;
Best Local Similarity 56.1%; Pred. No. 3.4e-23;
Matches 315; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

QY 2131 TCAGCAAGCGATTCCGACTCAGCGAGCGATTCCGACTCAGACAAATGACTCGGATTCAGAT 2190
Db TCACCTGGGACTTGACCTCAGCTGGAGATTTGGCCTCAGCTGAGTGTGACTTGACCTCAGCT 504
QY 2191 AGCGATTCTGACTCAGACAGTGTACTCAGATTCGACAGTGTACTCAGATTCAGATTCAGAT 2250
Db GGAGATTGGCCTCAGCTGAGTGTGGCTCAGCTGAGACTTGGCCTCAGCTGAGACTTGGCCT 444
QY 2251 TCTGACTCAGACAGTGTACTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 2310
Db TTGACTCAGCTGGAGATTTGGCCTCAGCTGGAGATTTGGCCTCAGCTGGAGATTTGGCT 384
QY 2311 TCCGACAGTGTATTCGCGACTCAGACAGCGATTCTGACTCCGACAGTGTATTCGCGACTCAGAC 2370
Db TCAGCTGGAGATTTGGCCTCAGATGTGACTTGGCCTCAGCTGAGACTTTCAGTGTAGCT 324
QY 2371 AGCGATTCTGATTCCGACAGTGTATTCGCGACTCAGATTCAGATTCAGATTCAGATTCAGAT 2430
Db GGAGATTGGCTTTCAGCTGTGCTGCTTGGCTTTCAGCTGTGCTGCTTGGCTTTCAGCTGTGCT 264

KEYWORDS EST.

Query Match 4.7%; Score 162.8; DB 7; Length 544;
Best Local Similarity 56.3%; Pred. No. 2.8e-22;
Matches 305; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

QY 2150 CACGGACGGATTCGGACTCAGACATGACTCGGATTCAGATTCGATTCGACTCAGACA 2209

```
Db 544 CAGCTGGAGATTGGCCCTCAGCTGGTGTGACTTGACCTCAGCTCAGTGGAGATTGGCCCTCAGCTG 485
QY 2210 GTGACTCAGATTCCGACAGTGAATCAGATTTCAGATAGCGATTCTGACTCAGACAGTGAAT 2269
Db 484 GTGACTTGGCTCAGCTGAGACTTGGCCCTCAGCTGGTGAATTTGGCTTTCAGCTCAGCTGGAGATT 425
QY 2270 CAGATTTCAGATAGCGATTTCAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 2329
Db 424 TGGCCCTCAGCTGGAGATTTCAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 365
QY 2330 CAGACAGCGATTTCAGCTCCGACAGTGAATTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 2389
Db 364 CAGATGGTGAATTCAGCTGGAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 305
QY 2390 GTGATTTCAGCTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 2449
Db 304 GTGACTTGGCTCAGCTGGAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 245
QY 2450 CAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 2509
Db 244 TGGCCCTCAGCTGGAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 185
QY 2510 CCGACAGTGAATTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 2569
Db 184 CACCTGGAGACTTCAGTGGAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 125
QY 2570 GTGACTCAGCTCAGACAGTGAATTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 2629
Db 124 GAGACTTGGCTCAGCTGGAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 65
QY 2630 CCGACTCCGACAGTGAATTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 2689
Db 64 TGGCCCTCAGCTGGAGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATTTCAGATAGCGATT 5
QY 2690 CA 2691
Db 4 CA 3
```

```
RESULT 14
CK423535
LOCUS CK423535
DEFINITION AUF Ipspn_68_h01 Spleen cDNA library Ictalurus punctatus cDNA 5',
mRNA sequence.
ACCESSION CK423535
VERSION CK423535.1 GI:40663235
KEYWORDS EST.
SOURCE Ictalurus punctatus (channel catfish)
ORGANISM Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
1 (bases 1 to 973)
Li, Z., Li, P., Liu, L., He, C., Kucuktas, H., Feng, J., Chen, L.,
Peatman, E., Baoprasertkul, P., Simmons, M., Muir, W., Grizzle, J.,
Dunham, R., and Brady, Y.
30,000 new catfish ESTs: new resources for functional analysis of
genes involved in aquaculture performance traits
Unpublished (2004)
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: T7.
FEATURES
1. 973 Location/Qualifiers
/organism="Ictalurus punctatus"
/mol_type="mRNA"
```

```
ORIGIN
Query Match 4.6%; Score 162.2; DB 7; Length 973;
Best Local Similarity 50.3%; Pred. No. 3.9e-22;
Matches 424; Conservative 0; Mismatches 415; Indels 4; Gaps 1;
QY 2084 GTGATTTCAGATTTCAGCAAGTGAATTCAGACTCAGCGAGTGAATTCAGATTTCAGCAAGCGATT 2143
Db 54 GAGAGAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 113
QY 2144 CCGACTCAGCGAGCGATTTCAGACTCAGACAATGACTCGGATTTCAGATAGCGATTTCGACT 2203
Db 114 GAGAGTGAAGTGAAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 173
QY 2204 CAGACAGTGAATTCAGATTTCAGACAGTGAATTCAGATTTCAGATAGCGATTTCGACTCAGACA 2263
Db 174 GAGATGGAGAGAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 233
QY 2264 GTGACTCAGATTTCAGATAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAAT 2323
Db 234 TGGAGAGAGAGACATAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 293
QY 2324 CCGACTCAGACAGCGATTTCGACTCCGACAGTGAATTCGACTCAGACAGCGATTTCAGATT 2383
Db 234 GAGAGAGAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 353
QY 2384 CCGACAGTGAATTCGACTCAGATAGCGATTTCGACTCAGATAGCGATTTCAGATTTCAGACA 2443
Db 354 GAGAGAGACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 413
QY 2444 CGGATTTCAGATTTCAGACAGCGATTTCAGATTTCAGATAGCGATTTCAGATTTCGCAAGTGAAT 2503
Db 414 GACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473
QY 2504 CAGATTTCGACAGTGAATTCGACTCCGATTTCAGATTTCGACTCAGACTCAGATT 2563
Db 474 GAGAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 533
QY 2564 CCGACAGTGAATTCAGACTCAGACAGTGAATTCGATTTCAGCGAGTGAATTCGAGATTTCAGATA 2623
Db 534 GATAGAGAGACAGAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 593
QY 2624 GTGATTTCGACTCCGACAGTGAATTCGATTTCAGATAGCGACTCAGACTCAGGATAGCGACT 2683
Db 594 GAGAGACACATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 653
QY 2684 CCGATTTCAGATAGCGATTTCGACTCCGACTCAGATAGCGATTTCAGATTTCAGATTTCAGATTTC 2739
Db 654 GAGACNGAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 713
QY 2740 GAATTCAGACAGCGATTTCAGATTTCAGACAGCGACTCAGACAGTGAATTCAGATTTCAGATTTCAG 2799
Db 714 GAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 773
QY 2800 GACTCGGATTTCAGCGAGTGAATTCAGACTCAGTAGTGAATTCGATTTCAGATTTCAGATTTCAG 2859
Db 774 CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833
QY 2860 GACTCAGAAAGTGAATTCAGATTTCAGATTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 2919
Db 834 GTGACAGACAGAGAGTGAATTCAGATTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 893
QY 2920 CCT 2922
Db 894 CCT 896
```

RESULT 15
ALI62088/c

LOCUS AL162088 1526 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp761l11023_r1.761 (synonym: hamy2) Homo sapiens cDNA clone
DKFZp761l11023_5', mRNA sequence.
ACCESSION AL162088
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1526)
AUTHORS Bloecker,H., Boechar,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GfR (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
SI sequence also available.
This clone (DKFZp761l1023) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..1526
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp761l1023"
/tissue_type="amygdala"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="761 (synonym: hamy2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 4.6%; Score 160.2; DB 1; Length 1526;
Best Local Similarity 62.4%; Pred. No. 1e-21;
Matches 285; Conservative 0; Mismatches 168; Indels 4; Gaps 2;
QY 2354 GTGATTCGGACTCAGACGCGATTCCGACAGTGATTCGGACTCAGATGCGGATT 2413
DB 1526 GAGATAGACACACAGATAGAGATAGATAGATAGATAGAT--AGATAGAGATAGAGATA 1469
QY 2414 CCGACTCAGATAGCGACTCAGATTCCAGACGCGATTCCAGATTCAGACAGCGGATTCAGATT 2473
DB 1468 GAGACACAGATAGAGATAGAGATAGATAGATAGATAGATAGATAGATAGATAGACACA 1409
QY 2474 CAGATAGCGGATTCAGATTCGG--ACAGTGACTCAGATTCGGACAGTGACTCGGATTCAGA 2531
DB 1408 CAGATAGAGATAGAGATAGAGATATAGATAGAGACACAGATAGATAGATAGATAGACA 1349
QY 2532 TAGCGATTTCAGATTCGACAGTGACTCAGATTCCGACAGTGACTCAGACTCAGACAGTGGA 2591
DB 1348 TAGAGATAGACACACAGATAGAGATAGAGATAGAGATAGAGATAGAGATAGAGACACAGA 1289
QY 2592 TTCGATTTCAGCGGATTCGGATTTCAGATAGTGATTCCGACTCCGACAGTGACTCGGA 2651
DB 1288 TAGAGATAGAGATAGAGATAGAGACACAGATAGAGATAGATAGATAGATAGAGACACAGA 1229
QY 2652 TTCAGATAGCGACTCAGACTCGGATTCAGATAGCGACTCGGATTCAGATAGCGGACTCAGA 2711
DB 1228 TAGAGATAGAGATAGATAGAGATAGAGATAGAGATAGAGATAGATAGATAGACACAGA 1169
QY 2712 TAGCGATTTCAGAAATTCAGACGCGGATTTCAGAAATTCAGACAGCGGATTTCAGATTTCAGACGCGGA 2771
DB 1168 TAGAGATAGAGATAGAGATAGAGATAGAGATAGAGATAGAGATAGAGATAGAGATAGAGATA 1109

QY 2772 CTCAGACAGTGACTCAGATTTCAGATAGTAGTGAATTCGGAT 2808
DB 1108 TAGAGATAGAGACACAGATAGATAGATAGATAGATAGAT 1072

Search completed: November 5, 2004, 12:54:47
Job time : 10227 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:48:10 ; Search time 176.713 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2

Perfect score: 4725

Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSLLFRKKENKKK 933

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_23Sep04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4725	100.0	933	3 AAY58435	Aay58435 Staphyloc
2	4725	100.0	933	4 AAB69508	Aab69508 Staphyloc
3	4717	99.8	933	6 ABJ18947	Abj18947 Pathogen
4	4662	98.7	927	6 ABM72221	Abm72221 Staphyloc
5	4662	98.7	936	2 AAW89801	Aaw89801 Staphyloc
6	4359	92.3	935	6 ABU16402	Abu16402 Protein e
7	4357	92.2	1021	4 AAU33975	Aau33975 Staphyloc
8	4357	92.2	1021	4 AAU36951	Aau36951 Staphyloc
9	2671	56.5	520	5 AAE29262	Aae29262 Staphyloc
10	2514	53.2	496	5 AAU75490	Aau75490 S. aureus
11	2514	53.2	496	6 ADA89664	Ada89664 Staphyloc
12	2514	53.2	496	6 ADA89663	Ada89663 Staphyloc
13	1813.5	38.4	913	6 ABJ18917	Abj18917 Pathogen
14	1808.5	38.3	918	2 AAY08640	Aay08640 S. aureus
15	1727	36.6	331	5 AAE29263	Aae29263 Staphyloc
16	1727	36.6	345	2 AAW31555	Aaw31555 Fibronect
17	1654.5	35.0	877	6 ADA89539	Ada89539 Staphyloc
18	1654.5	35.0	877	6 ABM72702	Abm72702 Staphyloc
19	1638.5	34.7	877	6 ABU42504	Abu42504 Protein e
20	1628	34.5	1092	2 AAW41602	Aaw41602 Staphyloc
21	1628	34.5	1092	7 ABM79019	Abm79019 Staphyloc
22	1607.5	34.0	1633	6 ABU42513	Abu42513 Protein e
23	1584.5	33.5	1802	3 AAY83170	Aay83170 Cell wall
24	1584.5	33.5	1802	3 AAY70119	Aay70119 Staph. ep
25	1552	32.8	1920	6 ABU43489	Abu43489 Protein e

26	1509	31.9	743	6 ADA89690	Ada89690 Staphyloc
27	1508.5	31.9	995	6 ABM72437	Abm72437 Staphyloc
28	1414	29.9	1166	2 AAY08643	Aay08643 S. aureus
29	1414	29.9	1166	6 ABJ18982	Abj18982 Pathogen
30	1402	29.7	717	6 ABJ18912	Abj18912 Pathogen
31	1398	29.6	513	6 ABU31230	Abu31230 Protein e
32	1395.5	29.5	1385	6 ABU16400	Abu16400 Protein e
33	1386.5	29.3	670	6 ABU42520	Abu42520 Protein e
34	1356	28.7	1637	6 ABJ19113	Abj19113 Pathogen
35	1354.5	28.7	953	6 ABU16533	Abu16533 Protein e
36	1320.5	27.9	1141	6 ABU42327	Abu42327 Protein e
37	1297.5	27.5	947	6 ABJ18940	Abj18940 Pathogen
38	1281	27.1	932	4 AAU36845	Aau36845 Staphyloc
39	1281	27.1	932	4 AAU34082	Aau34082 Staphyloc
40	1265.5	26.8	1349	4 AAU37544	Aau37544 Staphyloc
41	1265.5	26.8	1349	4 AAU34402	Aau34402 Staphyloc
42	1265.5	26.8	1349	6 ABM72436	Abm72436 Staphyloc
43	1249.5	26.4	930	2 AAY08641	Aay08641 S. aureus
44	1131	23.9	275	6 ADA89579	Ada89579 Staphyloc
45	1110.5	23.5	1315	2 AAY08642	Aay08642 S. aureus

ALIGNMENTS

RESULT 1
AAY58435
ID AAY58435 standard; protein; 933 AA.
XX AC
XX AAY58435;
XX 27-MAR-2000 (first entry)
XX Staphylococcus aureus fibrinogen binding ClfA protein.
XX ClfA; fibrinogen binding protein; bacterial colonisation;
XX KW indwelling medical device; staphylococcal infection.
XX KW
XX OS Staphylococcus aureus.
XX

Key Location/Qualifiers
FT Peptide 1..39
FT /note= "Signal peptide"
FT Region 40..559
FT /note= "Region A"
FT Region 332..550
FT /note= "Fibrinogen-binding region"
FT Region 560..867
FT /note= "Region R"
FT Region 896..900
FT /note= "Gram positive wall-associated consensus motif"

US6008341-A.

28-DEC-1999.

22-AUG-1994; 94US-00293728.

22-AUG-1994; 94US-00293728.

(QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.

Foster TJ, Mcdevitt DL;

WPI; 2000-096389/08.

N-PSDB; AAZ55832.

Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus aureus, useful for treatment or prevention of infections.

Disclosure; Fig 2A-1-4; 35pp; English.

This sequence represents the Staphylococcus aureus fibrinogen-binding

app. cat

QY 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFLLSSKEADASENSVTQSDASNEKSDSS 60
DB 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFLLSSKEADASENSVTQSDASNEKSDSS 60
QY 61 SVSAAPKTDITVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
DB 61 SVSAAPKTDITVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
QY 121 TTNQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSTNAENVTQDST 180
DB 121 TTNQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSTNAENVTQDST 180
QY 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240
DB 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240
QY 241 VGDSTGTTVPHQAGYVKNLNGFVSVPNSAVKGTFTKIVPEKELNGLVGTSTAKVPPIMAG 300
DB 241 VGDSTGTTVPHQAGYVKNLNGFVSVPNSAVKGTFTKIVPEKELNGLVGTSTAKVPPIMAG 300
QY 301 DQVLANGVIDSGNVIYFTDYVNTKDDVKAATLMPAYIDPENVKKTGNVTLATIGSTT 360
DB 301 DQVLANGVIDSGNVIYFTDYVNTKDDVKAATLMPAYIDPENVKKTGNVTLATIGSTT 360
QY 361 ANKTVLVDYKGYKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAPLVGNLKPNT 420
DB 361 ANKTVLVDYKGYKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAPLVGNLKPNT 420
QY 421 DSNALIDQNTSIKVKYVDNAADLSESYFVNPENFEDVNTSVNITFPNPQYKVEFNTPD 480
DB 421 DSNALIDQNTSIKVKYVDNAADLSESYFVNPENFEDVNTSVNITFPNPQYKVEFNTPD 480
QY 481 DQITTPYIVVNGHIDPNSGDLALRSTLYGNSNIIWRMSWDNEVAFNNGSGDGDID 540
DB 481 DQITTPYIVVNGHIDPNSGDLALRSTLYGNSNIIWRMSWDNEVAFNNGSGDGDID 540
QY 541 KPVPPEQDPGELEPIPEDSDPGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSD 600
DB 541 KPVPPEQDPGELEPIPEDSDPGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSD 600
QY 601 SDSDSASDSDSASD 660
DB 601 SDSDSASDSDSASD 660
QY 661 SD 720
DB 661 SD 720
QY 721 SD 780
DB 721 SD 780
QY 781 SD 840
DB 781 SD 840
QY 841 SD 900
DB 841 SD 900
QY 901 SEDEANTSLIWGLLASIGSLLLFRRKKENKDKK 933
DB 901 SEDEANTSLIWGLLASIGSLLLFRRKKENKDKK 933

RESULT 3
ID ABJ18947
XX ABJ18947 standard; protein; 933 AA.
AC ABJ18947;
XX
DT 06-MAR-2003 (first entry)
XX

DE Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW auto-immune disease; HIV; hepatitis.
XX Staphylococcus sp.
XX WO200259148-A2.
XX 01-AUG-2002.
XX 21-JAN-2002; 2002WO-EP000546.
XX 26-JAN-2001; 2001AT-00000130.
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX Meinke A, Nagy E, Von Ahlsen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;
PI Tempelmaier B;
DR WPI; 2003-075410/07.
XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.
XX Example 7; Page 160; 252pp; English.
XX The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX allergen, a tissue or host prone to auto-immunity, where the antigens are
XX used in a vaccine, comprises providing antibody preparation from a plasma
XX pool of a type of animal, or individual sera with antibodies against the
XX specific pathogen, tumour, allergen, tissue or host prone to auto-
XX immunity. The hyperimmune serum-reactive antigens comprising any of the
XX 62 sequences of 53-2261 amino acids fully defined in the specification,
XX or their hyperimmune fragments are useful for the manufacture of a
XX pharmaceutical preparation, particularly a vaccine against staphylococcal
XX infections or colonisation against S. aureus or S. epidermidis. The
XX preparation of antibodies is useful for the manufacture of a medicament
XX for treating or preventing staphylococcal infections or colonisation
XX against S. aureus or S. epidermidis. The antibody preparations may also
XX be used for diagnostic and imaging purposes. Other conditions that can be
XX treated include cancer, autoimmune diseases or infections caused by viral
XX (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX sequence represents a staphylococcal protein relating to the method for
XX identifying and producing pathogen specific antigens of the invention
XX
SQ Sequence 933 AA;
Query Match 99.8%; Score 4717; DB 6; Length 933;
Best Local Similarity 99.9%; Pred. No. 1.5e-232;
Matches 932; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFLLSSKEADASENSVTQSDASNEKSDSS 60
DB 1 MNMKEKHAIRKKSIGVASVLVGTLLIGFLLSSKEADASENSVTQSDASNEKSDSS 60
QY 61 SVSAAPKTDITVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
DB 61 SVSAAPKTDITVSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEPTVTEATT 120
QY 121 TTNQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSTNAENVTQDST 180
DB 121 TTNQANTPATQSSNTNAEELVNOTSNFTFNDTNTVSSVNSPQNSTNAENVTQDST 180
QY 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240
DB 181 EATPSNNEAPQSTDSAKDQVNVQAVNTSAPRMAFSLAAVAADAPAGTDITNQLNVT 240

QY 241 VGIDSGTTVPYHQAGYVVKLYGFSVPNSAVKGDFTKIVPKELNLNGVTSTAKVPPIMAG 300
DB 241 VGIDSGTTVPYHQAGYVVKLYGFSVPNSAVKGDFTKIVPKELNLNGVTSTAKVPPIMAG 300
QY 301 DOVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
DB 301 DOVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
QY 361 ANKTVLDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKENT 420
DB 361 ANKTVLDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKENT 420
QY 421 DSNALIDQONTSIKIVYKVDNAADLSSEYFVNPENFEDVTNSVNIIFPNPNQYKVFENTPD 480
DB 421 DSNALIDQONTSIKIVYKVDNAADLSSEYFVNPENFEDVTNSVNIIFPNPNQYKVFENTPD 480
QY 481 DOITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNIIWRSMGWDNEVAFNNGSGGID 540
DB 481 DOITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNIIWRSMGWDNEVAFNNGSGGID 540
QY 541 KPVVPEQDPDEGEIEPIPEDSDPGSDSGSDNSDGSSTSDSGSDSASDSASA 600
DB 541 KPVVPEQDPDEGEIEPIPEDSDPGSDSGSDNSDGSSTSDSGSDSASDSASA 600
QY 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 660
DB 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 660
QY 661 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 720
DB 661 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 720
QY 721 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 780
DB 721 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 780
QY 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 840
DB 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 840
QY 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 900
DB 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 900
QY 901 SEDEANTSLIWGLLASIGSLLLFRKKENKOKK 933
DB 901 SEDEANTSLIWGLLASIGSLLLFRKKENKOKK 933

RESULT 4

ID ABM72221
XX ABM72221 standard; protein; 927 AA.
AC ABM72221;
XX 20-NOV-2003 (first entry)
XX Staphylococcus aureus protein #1461.
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target.
XX Staphylococcus aureus.
OS WO2002094868-A2.
XX 28-NOV-2002.
XX 27-MAR-2002; 2002WO-IB002637.
XX 27-MAR-2001; 2001GB-00007661.
XX

PA (CHIR-) CHIRON SPA.
XX Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
DR N-PSDB; ACF73781.
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX Claim 1; SEQ ID NO 2922; 49pp; English.
XX The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX Sequence 927 AA;
XX
Query Match 98.7%; Score 4662; DB 6; Length 927;
Best Local Similarity 99.0%; Pred. No. 9.4e-230;
Matches 924; Conservative 0; Mismatches 3; Indels 6; Gaps 1;
QY 1 MMKKKEKHAIKKSIGVASVLVGLTIGFLLSSKEADASENSVTQSDASNESKSDSS 60
DB 1 MMKKKEKHAIKKSIGVASVLVGLTIGFLLSSKEADASENSVTQSDASNESKSDSS 60
QY 61 SVSAAPKTDITNVSDTKTSSNTNGETSVAQNPQOETTQSSSTNATTEETPTVTEATT 120
DB 61 SVSAAPKTDITNVSDTKTSSNTNGETSVAQNPQOETTQSSSTNATTEETPTVTEATT 120
QY 121 TTQANTPATTQSSNTNAEELVNQTSNETTNDTNTVSSVNSPQNSTNAENVSTTQDTST 180
DB 121 TTQANTPATTQSSNTNAEELVNQTSNETTNDTNTVSSVNSPQNSTNAENVSTTQDTST 180
QY 181 EATPSNESAPQSDTASNDKVVQAVNTSAPRMRAFSLAAVAADAPAACTITNLTNTV 240
DB 181 EATPSNESAPQSDTASNDKVVQAVNTSAPRMRAFSLAAVAADAPAACTITNLTNTV 240
QY 241 VGIDSGTTVPYHQAGYVVKLYGFSVPNSAVKGDFTKIVPKELNLNGVTSTAKVPPIMAG 300
DB 241 VGIDSGTTVPYHQAGYVVKLYGFSVPNSAVKGDFTKIVPKELNLNGVTSTAKVPPIMAG 300
QY 301 DOVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
DB 301 DOVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
QY 361 ANKTVLDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKENT 420
DB 361 ANKTVLDYKGYKGFYNLSIKGTIDQIDKTNNTYRQTIYVNPFGDNVIAPIVLTGNLKENT 420
QY 421 DSNALIDQONTSIKIVYKVDNAADLSSEYFVNPENFEDVTNSVNIIFPNPNQYKVFENTPD 480
DB 421 DSNALIDQONTSIKIVYKVDNAADLSSEYFVNPENFEDVTNSVNIIFPNPNQYKVFENTPD 480
QY 481 DOITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNIIWRSMGWDNEVAFNNGSGGID 540
DB 481 DOITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNIIWRSMGWDNEVAFNNGSGGID 540
QY 541 KPVVPEQDPDEGEIEPIPEDSDPGSDSGSDNSDGSSTSDSGSDSASDSASA 600
DB 541 KPVVPEQDPDEGEIEPIPEDSDPGSDSGSDNSDGSSTSDSGSDSASDSASA 600
QY 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 660
DB 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 660

CC DNA sequences can be used for recombinant production of the polypeptides.
CC The new DNA sequences (and their fragments) are useful as primers or
CC probes for isolating homologues of any of the 5191 S.aureus DNA sequences
CC contained on the computer readable medium
XX
SQ Sequence 936 AA;

Query Match 98.7%; Score 4662; DB 2; Length 936;
Best Local Similarity 99.0%; Pred. No. 9.5e-230;
Matches 924; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 1 MNMKKEKHAIRKKSIGVASVLVGTLLIGFLLSKKADASENSVTQSDSASNSKNSDSS 60
DB 10 MNMKKEKHAIRKKSIGVASVLVGTLLIGFLLSKKADASENSVTQSDSASNSKNSDSS 69
QY 61 SVSAAKPTDDTNTVSDTKTSNTNNGTSTVAONPAQOETTCOSSSTNAETTEPVTGEATTT 120
DB 70 SVSAAKPTDDTNTVSDTKTSNTNNGTSTVAONPAQOETTCOSSSTNAETTEPVTGEATTT 129
QY 121 TTNQANTPATTSQNTNNAEELVNQTSNETTFNDTNTVSSVNSPQNTNAENVTQDTST 180
DB 130 TTNQANTPATTSQNTNNAEELVNQTSNETTFNDTNTVSSVNSPQNTNAENVTQDTST 189
QY 181 EATPSNNESAPQSDASNDKDVNQAVNTSAPRMRAPSLAAVAADAPAGTDTITNQLTNTV 240
DB 190 EATPSNNESAPQSDASNDKDVNQAVNTSAPRMRAPSLAAVAADAPAGTDTITNQLTNTV 249
QY 241 VGIDSGTIVYVPHQAGYKLVNYSFVPSNAVKGTGTFKITVPELNLNGVSTAKVPPIMAG 300
DB 250 VGIDSGTIVYVPHQAGYKLVNYSFVPSNAVKGTGTFKITVPELNLNGVSTAKVPPIMAG 309
QY 301 DOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYIDIPENVKKTGNVLTATGISTT 360
DB 310 DOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYIDIPENVKKTGNVLTATGISTT 369
QY 361 ANKTVLVDEYKYGKFNLSIKGTIDQDKNNVTYRQTIYVNPSSGDNVIAPIVLTGNLKPNT 420
DB 370 ANKTVLVDEYKYGKFNLSIKGTIDQDKNNVTYRQTIYVNPSSGDNVIAPIVLTGNLKPNT 429
QY 421 DSNALIDQOQNTSIKVVYKVDNAADLSESYFVNPENFEDVTNSVNIITFPNPKYKVEEPTD 480
DB 430 DSNALIDQOQNTSIKVVYKVDNAADLSESYFVNPENFEDVTNSVNIITFPNPKYKVEEPTD 489
QY 481 DQITTPYIVVNGHIDPNSKGLALRSTLYGNSNIIWRMSWMDNEVAFNNGSGSGDID 540
DB 490 DQITTPYIVVNGHIDPNSKGLALRSTLYGNSNIIWRMSWMDNEVAFNNGSGSGDID 549
QY 541 KPVVPEQDPGEIETPIEDSDSDPGSDSGSDNSGSDSGSDSTSDSGSDSASDSDSA 600
DB 550 KPVVPEQDPGEIETPIEDSDSDPGSDSGSDNSGSDSGSDSTSDSGSDSASDSDSA 609
QY 601 SDSDSASDSDSASD 660
DB 610 SDSDSASDSDSASD 669
QY 661 SD 720
DB 670 SD 729
QY 721 SD 780
DB 730 SD 789
QY 781 SD 840
DB 790 SD 843
QY 841 SD 900
DB 844 SD 903
QY 901 SEDEANTSLIWGLIASIGSLLLFRKKENKDKK 933
DB 933 SEDEANTSLIWGLIASIGSLLLFRKKENKDKK 933

QY 661 SD 720
DB 661 SD 720
QY 721 SD 780
DB 721 SD 780
QY 781 SD 840
DB 781 SD 834
QY 841 SD 900
DB 835 SD 894
QY 901 SEDEANTSLIWGLIASIGSLLLFRKKENKDKK 933
DB 895 SEDEANTSLIWGLIASIGSLLLFRKKENKDKK 927

RESULT 5
AAW89801
ID AAW89801 standard; protein; 936 AA.
XX AAW89801;
AC AAW89801;
DT 16-MAR-1999 (first entry)
XX Staphylococcus aureus protein SEQ ID #5249.
DE
XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome.
XX Staphylococcus aureus.
OS
XX EF786519-A2.
XX 30-JUL-1997.
XX 07-JAN-1997; 97EP-00100117.
XX 05-JAN-1996; 96US-0009861P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kumsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX WPI; 1997-374922/35.
XX
XX polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX stored on computer readable medium and used in the production of anti-
XX S.aureus vaccines.
XX
XX Claim 23; Page 3255-3258; 3271pp; English.

This sequence represents a Staphylococcus aureus protein sequence of the invention. The DNA sequences encoding the S.aureus proteins are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S.aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S.aureus infection. The polypeptides can also be used in a kit for the immunodetection of S.aureus in a sample. S.aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the

Db 904 SEDEANTSLIWGLLASIGSLLFRKKNKDKK 936

RESULT 6

ABU16402
ID ABU16402 standard; protein; 935 AA.

AC ABU16402;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #1929.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Staphylococcus aureus.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 23-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen Ku, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA20272.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 44326; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 935 AA;

SQ

Query Match

Best Local Similarity 92.3%; Score 4359; DB 6; Length 935;

Matches 865; Conservative 28; Mismatches 40; Indels 2; Gaps 1;

QY 1 MMKKKEKHAIRKKSIGVASVLVGTLLIGFGLSSKEADASENSVTOSDSASNEKNDSS 60

Db 1 MMKKKEKHAIRKKSIGVASVLVGTLLIGFGLSSKEADASENSVTOSDSASNEKNDSS 60

QY 61 SVSAAFKTDDTVNSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGATTT 120

Db 61 SVSAAFKTDDTVNSDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGATTT 120

QY 121 TTNQANTPATTOSSNTNABELVQNSNETTFNDTNTVSVNSPQNSNAENVSTTQDTST 180

Db 121 TTNQANTPATTOSSNTNABELVQNSNETTFNDTNTVSVNSPQNSNAENVSTTQDTST 180

QY 181 EATPSNNEAPOSTDASNKDVVNOAVNTSAPWRAPSLAAVAADAPAACTDITNOLNTVT 240

Db 181 EATPSNNEAPONTDASNKDVVNOAVNTSAPWRAPSLAAVAADAPAACTDITNOLNTVT 240

QY 241 VGIDSGTTVYPHQAGYVKLYGFSVENSAAKGDTFKITVPKELNLNGVTSTAKVPPIMAG 300

Db 241 VTIDSGTTVYPHQAGYVKLYGFSVENSAAKGDTFKITVPKELNLNGVTSTAKVPPIMAG 300

QY 301 DOVLANGVIDSGNVIIYFTDYVNTKDDVKALTMPAYIDPENVKKTGNVTATGIGSTT 360

Db 301 DOVLANGVIDSGNVIIYFTDYVNTKDDVKALTMPAYIDPENVKKTGNVTATGIGSTT 360

QY 361 ANKTVLVDEYKGYKFNLSIKGTIIDIQIKNTNTYQTIYVNSGDNVAPLVITGNLKPT 420

Db 361 ASKTVLVDEYKGYKFNLSIKGTIIDIQIKNTNTYQTIYVNSGDNVAPLVITGNLKPT 420

QY 421 DSNALIDQONTSIKVKVDNAADLSSEYFVNPFEDVTNSVNIITFPNPNQYKVFNTPD 480

Db 421 KSNALIDAKNTDIKVYVDNANDLSSEYFVNPFEDVTNQVRIFFPNANQYKVFNTPD 480

QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGNSNIIWRSMWNEVAFNNGSGSGDGD 540

Db 481 DOITTPYIVVNGHIDPASTGDLALRSTFYGYDSNFIWRSMWNEVAFNNGSGSGDGD 540

QY 541 KPVVPEQDEPGEIPIPEDSDPGSDGSDNSDSDSDSDSDSDSDSDSDSDSDSDSDSA 600

Db 541 KPVVPEQDEPGEIPIPEDSDPGSDGSDNSDSDSDSDSDSDSDSDSDSDSDSDSDSA 600

QY 601 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 660

Db 601 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 660

QY 661 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 720

Db 661 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 720

QY 721 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 780

Db 721 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 780

QY 781 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 838

Db 781 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 840

QY 839 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 898

Db 841 SDSGSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASD 900

QY TGSDEANTSLIWGLLASIGSLLFRKKNKDKK 933

Db TGSDEANTSLIWGLLASIGSLLFRKKNKDKK 935

[illegible]

DT 14-FEB-2002 (first entry)
 XX Staphylococcus aureus cellular proliferation protein #1121.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS54810.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 XX Example 3; SEQ ID NO 12544; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1021 AA;

Query Match 92.2%; Score 4357; DB 4; Length 1021;
 Best Local Similarity 85.5%; Pred. No. 3.8e-214;
 Matches 873; Conservative 27; Mismatches 33; Indels 88; Gaps 3;

QY 1 MNMKKEKHARKKSIGVASVLVGLTIGFLLSKSEADASENSVTQSDASNESKSDSS 60
 DB 1 MNMKKEKHARKKSIGVASVLVGLTIGFLLSKSEADASENSVTQSDASNESKSDSS 60
 QY 61 SVSAAPKTDITNVSDTKTSNTNNGETSVQAQNPQOETTQSSSTNATTEPTVGEATT 120
 DB 61 SVNAAPKTDITNVSDTKTSNTNNGETSVQAQNPQOETTQSSSTNATTEPTVGEATT 120
 QY 121 TTQANTPATQSSNTNAELVNQTSNETTFNDNTVSSVNSPQNSTNAENVSTQDTST 180
 DB 121 ATKQANTPATQSSNTNAELVNQTSNETASNDINTVSSVNSPQNSTNAENVSTQDTST 180

QY 181 EATPSNNEAPQSTDSASNDKDVVNOAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTNVT 240
 DB 181 EATPSNNEAPQSTDSASNDKDVVNOAVNTSAPRMRAFSLAAVAADAPAGTDITNQLTDVK 240
 QY 241 VGIDSGTTVYVPHQAGYVVKLYNGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPPIMAG 300
 DB 241 VTIDSGTTVYVPHQAGYVVKLYNGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPPIMVG 300
 QY 301 DOVLANGVIDSDGNVIYTFDYNTKDYKATILTMAYIDPENVKITGNVLTATGSGTT 360
 DB 301 DOVLANGVIDSDGNVIYTFDYNTKDYKATILTMAYIDPENVKITGNVLTATGSGTT 360
 QY 361 ANKTVLDYKCYKGFYNLSIKGTIDDKTNNTYRQTIYVNPDSGDNVIAFVLTKGNLKPN 420
 DB 361 ASKTVLDYKCYKGFYNLSIKGTIDDKTNNTYRQTIYVNPDSGDNVIAFVLTKGNLKPN 420
 QY 421 DSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTNSVNIITPPNQYKVEFTED 480
 DB 421 KGNALIDAKNTDIKVKYVDNANDLSESYFVNPENFEDVTNQVRISFPNANQYKVEFTDD 480
 QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYNSNIIFWESMWDNEVAFNNGSGSDGID 540
 DB 481 DOITTPYIVVNGHIDPASTGDLALRSTFYGYDSNFIWESMWDNEVAFNNGSGSDGID 540
 QY 541 KFWPEQPDPEGEIPIPEDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
 DB 541 KFWPEQPDPEGEIPIPEDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
 QY 601 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 622
 DB 601 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 660
 QY 623 SNDSD 682
 DB 661 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 720
 QY 683 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 742
 DB 721 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 780
 QY 743 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 802
 DB 781 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 840
 QY 803 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 832
 DB 841 SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 900
 QY 833 --SDDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 872
 DB 901 SGSDSDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 960
 QY 873 PPNPFGKGTNASNKNEAKDSKEPLPDTGSEDEANTSLIWGLLASIGSLLLFRKKEKDK 932
 DB 961 PPNPFGKGTNASNKNEAKDSKEPLPDTGSEDEANTSLIWGLLASIGSLLLFRKKEKDK 1020
 QY 933 K 933
 DB 1021 K 1021

RESULT 9
 AAE29262
 ID AAE29262 standard; protein; 520 AA.
 XX
 AC AAE29262;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Staphylococcus aureus Clf40 protein.
 XX
 KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
 immunological; staphylococcal infection; impetigo; pneumonia; furuncle;

KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

XX Staphylococcus aureus.

OS

XX

PN WO200272600-A2.

XX

PD 19-SEP-2002.

XX

PF 28-JAN-2002; 2002WO-US002296.

XX

PR 26-JAN-2001; 2001US-0264072P.

PR 12-MAR-2001; 2001US-0274611P.

PR 18-JUN-2001; 2001US-0298413P.

PR 30-JUL-2001; 2001US-0308116P.

XX

PA (INH1-) INHIBITEX INC.

XX

PI Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;

XX

PI WPI; 2002-759834/82.

XX

DR N-PSDB; AAD46861.

DR

XX

XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for

PT treating or preventing Staphylococcus aureus infection e.g. wound

PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

PT a human or animal.

XX

XX Claim 9; Page 67-69; 80pp; English.

PS

XX

XX The invention relates to monoclonal antibody which binds the clumping

CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA

CC monoclonal antibody is useful for treating or preventing S. aureus

CC infection in a human or animal, and for inhibiting the binding of

CC Staphylococcus aureus to fibrinogen or fibrin. The immunogenic fragment

CC of S. aureus ClfA protein, S. aureus Clf33 protein, or the S. aureus N3

CC protein is useful for inducing an immunological response in a human or

CC animal. These staphylococcal infections include wound infections, sepsis,

CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The

CC present sequence is Staphylococcus aureus Clf40 protein

XX

SQ Sequence 520 AA;

Query Match 56.5%; Score 2671; DB 5; Length 520;

Best Local Similarity 100.0%; Pred. No. 1.9e-128; Indels 0; Gaps 0;

Matches 520; Conservative 0; Mismatches 0;

QY 40 SENVTSQSDSASNESKNDSSVSAAPKTDITNNVSDTKTSSNTNGETSVAGNPAQQT 99

DB 1 SENVTSQSDSASNESKNDSSVSAAPKTDITNNVSDTKTSSNTNGETSVAGNPAQQT 60

QY 100 QSSSTNATTEPVTGEATTTTNOANTPATTSQNTNNAEELVNOTSNETTNDNTVSS 159

DB 61 QSSSTNATTEPVTGEATTTTNOANTPATTSQNTNNAEELVNOTSNETTNDNTVSS 120

QY 160 VNSPQNSTNAENVSTQDTSSTATSNESAPQSDASKNVNOAVNTSAPRMAFSLA 219

DB 121 VNSPQNSTNAENVSTQDTSSTATSNESAPQSDASKNVNOAVNTSAPRMAFSLA 180

QY 220 AVAADAAPAGTDTNLTNTVTVGIDSGTTVYVPHQAGYVKNLYGFSVPNSAVKGDTPKIIV 279

DB 181 AVAADAAPAGTDTNLTNTVTVGIDSGTTVYVPHQAGYVKNLYGFSVPNSAVKGDTPKIIV 240

QY 280 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVITFTDYVNTKDDVKATLTMPAYI 339

DB 241 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVITFTDYVNTKDDVKATLTMPAYI 300

QY 340 DENVKKTGNVLTATIGSTTANKTVLVDEYKYGKYNLSIKGTIDIDKTNNTVQTIV 399

DB 301 DENVKKTGNVLTATIGSTTANKTVLVDEYKYGKYNLSIKGTIDIDKTNNTVQTIV 360

QY 400 VNPSGDNVAPVLTGNLKNPTDSNALIDQNTSIKVKVDNADLSGESYFVNPFEDVT 459

DB 361 VNPSGDNVAPVLTGNLKNPTDSNALIDQNTSIKVKVDNADLSGESYFVNPFEDVT 420

QY 460 NSVNIFFPNPNQKVEFNTPDDOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIWR 519

DB 421 NSVNIFFPNPNQKVEFNTPDDOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIWR 480

QY 520 SMSWDNEVAFNNGSGSGDIDKPVVPEQDPDEGEIPIEPIE 559

DB 481 SMSWDNEVAFNNGSGSGDIDKPVVPEQDPDEGEIPIEPIE 520

RESULT 10

AAU75490

ID AAU75490 standard; protein; 496 AA.

XX

XX AC AAU75490;

XX

DT 23-APR-2002 (first entry)

DE

DE S. aureus antigenic protein associated protein #10.

XX

XX Antigenic protein; vaccine; SEREX; antibacterial; antiinflammatory;

XX dermatological; antiulcer; tuberculostatic; immunosuppressive;

XX septicaemia; food poisoning; skin disorders; peritonitis; endocarditis;

XX tuberculosis; blood infection; sepsis; meningitis; pneumonia;

XX stomach ulcer; gonorrhea; necrotising fasciitis; impetigo;

XX Lyme's disease; gastro-enteritis; dysentery; shigellosis.

XX

OS Staphylococcus aureus.

XX

XX WO200198499-A1.

XX

XX 27-DEC-2001.

XX

XX 20-JUN-2001; 2001WO-GB002685.

XX

XX 20-JUN-2000; 2000GB-00014907.

XX

XX (UYSH-) UNIV SHEPPFIELD.

PA (BIOS-) BIOSYNEXUS INC.

XX

XX Foster S, McDowell P, Brummell K, Clarke S;

XX

XX WPI; 9A002-106544-14.

XX

XX Identifying antigenic polypeptides expressed by pathogenic organisms

PT e.g., Staphylococcus aureus during infection, by SEREX (serological

PT identification of antigens by recombinant expression cloning) techniques.

XX

XX Disclosure; Page 72-74; 85pp; English.

XX

XX The invention relates to a method for identifying antigenic polypeptides

CC expressed by pathogenic organisms e.g., Staphylococcus aureus during

CC infection, by SEREX (serological identification of antigens by

CC recombinant expression cloning) techniques. The method involves providing

CC a nucleic acid library encoding genes/partial gene sequences of

CC pathogenic organisms, transforming/transfecting the library into host

CC cells, contacting the polypeptides expressed by the genes with autologous

CC antisera derived from an animal infected with, or has been infected with

CC the pathogen and purifying the nucleic acid encoding the polypeptide or

CC partial polypeptide binding to the antisera. Also included are the

CC nucleic acids and polypeptides isolated by the method, vectors and

CC transformed cells expressing them, a vaccine comprising the polypeptide

CC and the production of monoclonal antibodies against the polypeptides. The

CC protein and vaccine are useful for immunising an animal (preferably

CC human) against a pathogenic microbe. The proteins and antibodies are

CC useful for manufacturing a medicament for treating Staphylococcus aureus-

CC associated septicaemia, food poisoning or skin disorders or

CC Staphylococcus epidermidis-associated septicaemia, sepsis, meningitis,

CC endocarditis, tuberculosis, blood infections, necrotising fasciitis, impetigo,

CC pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, The present

CC sequence is an S. aureus protein sequence associated with the antigenic

CC proteins of the invention. Note: The present sequence is included in the

CC sequence listing but is not mentioned anywhere else in the specification
XX
SQ Sequence 496 AA;

Query Match 53.2%; Score 2514; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.9e-120;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNMKEKHAIRKKSIGVASVLVGLIGFLLSSKEADASENSVTQDSASNEKSDSS 60
Db 1 MNMKEKHAIRKKSIGVASVLVGLIGFLLSSKEADASENSVTQDSASNEKSDSS 60

Qy 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVQNPAAQETQSSSTNATTEPTVGEATT 120
Db 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVQNPAAQETQSSSTNATTEPTVGEATT 120

Qy 121 TTNOANTPATVQSSNTNABELVNQTSNETTFNDTNTVSSVNSPQNSTNAENVSTQDTST 180
Db 121 TTNOANTPATVQSSNTNABELVNQTSNETTFNDTNTVSSVNSPQNSTNAENVSTQDTST 180

Qy 181 EATPSNNEAPQSTDSASNDKDVVQAVNTSAPRMEAFSLAAVAADAPAAAGTDITNQLTNVT 240
Db 181 EATPSNNEAPQSTDSASNDKDVVQAVNTSAPRMEAFSLAAVAADAPAAAGTDITNQLTNVT 240

Qy 241 VGDSGTTVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVTSKAKVPPIMAG 300
Db 241 VGDSGTTVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVTSKAKVPPIMAG 300

Qy 301 DQVLANGVIDSDGNVIYTFDYNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360
Db 301 DQVLANGVIDSDGNVIYTFDYNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360

Qy 361 ANKTVLDEYKYGKFNLSIKGPIQIDKTNNTYRQTIYVNPSSGDNVIAPLVTLGNLKNPT 420
Db 361 ANKTVLDEYKYGKFNLSIKGPIQIDKTNNTYRQTIYVNPSSGDNVIAPLVTLGNLKNPT 420

Qy 421 DSNALLDQNTSTKVKVQVNAADLSSEYFVNPENFEDVNSVNIITPNPNQYKVEFTPD 480
Db 421 DSNALLDQNTSTKVKVQVNAADLSSEYFVNPENFEDVNSVNIITPNPNQYKVEFTPD 480

Qy 481 DQITTFYIVVNGHID 496
Db 481 DQITTFYIVVNGHID 496

RESULT 11

ADA89664

ID ADA89664 standard; protein; 496 AA.

AC ADA89664;

XX 20-NOV-2003 (first entry)

DE Staphylococcus aureus antigenic protein #203.

antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
antitumor; immunostimulant; ophthalmological; pathogenic microbe;
bacteremia; septic shock; organ infection; skin infection;
bacterial basal colonisation; bacterial eye infection; septicaemia;
tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
gastro-enteritis; dysentery; shigellosis; skin disorder.

Staphylococcus aureus.

OS WO2003011899-A2.

XX 13-FEB-2003.

XX 02-AUG-2002; 2002WO-GB003606.

PR

XX 02-AUG-2001; 2001GB-00018925.

XX 09-JAN-2002; 2002GB-00000349.

PA (UYSH-) UNIV SHEFFIELD.

PA (BIOS-) BIOSYNEXUS INC.

PI Foster S, Mond J, Clarke S, McDowell P, Brummel K;

XX WPI; 2003-256434/25.

XX

New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
useful as a vaccine for immunizing humans against e.g. bacteremia, septic
shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
impetigo.

Claim 4; Page 161; 189pp; English.

The present invention describes an antigenic protein or its part, which
is for use as a vaccine. The antigenic protein is encoded by an isolated
DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
sequences (designated dnaS and dna SE, respectively; and which encodes a
protein expressed by a pathogenic organism. Also described: (1) a vaccine
composition comprising at least one antigenic protein; (2) a method of
immunising an animal against a disease or condition caused by a
pathogenic microbe by administering the antigenic protein or the vaccine;
(3) an antibody or its binding part obtainable by the method above; (4)
preparing a hybridoma cell line producing monoclonal antibodies; (5) a
hybridoma cell line produced by the method of (4); and (6) identifying
opsonic antigens expressed by a pathogenic microbe. The antigenic
proteins have antibacterial, neuroprotective, immunosuppressive,
antiinflammatory, antitumor, immunostimulant and ophthalmological
activities, and can be used in vaccines. The antigenic proteins or
vaccines can be used for immunising an animal (specifically a human)
against a disease or condition caused by a pathogenic microbe, e.g.
bacteremia, septic shock, organ infection, skin infection, bacterial
basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
bacteria-associated food poisoning, blood infections, peritonitis,
endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
strep throat, streptococcal-associated toxic shock, necrotising
fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
skin disorders, S. epidermidis-associated septicaemia, peritonitis or
endocarditis. The present sequence represents a S. aureus antigenic
protein sequence from the present invention.

XX Sequence 496 AA;

SQ

Query Match

Best Local Similarity 53.2%; Score 2514; DB 6; Length 496;

Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 MNMKEKHAIRKKSIGVASVLVGLIGFLLSSKEADASENSVTQDSASNEKSDSS 60

Db 1 MNMKEKHAIRKKSIGVASVLVGLIGFLLSSKEADASENSVTQDSASNEKSDSS 60

Qy

61 SVSAAPKTDITNVSDTKTSSNTNNGETSVQNPAAQETQSSSTNATTEPTVGEATT 120

Db 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVQNPAAQETQSSSTNATTEPTVGEATT 120

Qy

121 TTNOANTPATVQSSNTNABELVNQTSNETTFNDTNTVSSVNSPQNSTNAENVSTQDTST 180

Db 121 TTNOANTPATVQSSNTNABELVNQTSNETTFNDTNTVSSVNSPQNSTNAENVSTQDTST 180

Qy

181 EATPSNNEAPQSTDSASNDKDVVQAVNTSAPRMEAFSLAAVAADAPAAAGTDITNQLTNVT 240

Db 181 EATPSNNEAPQSTDSASNDKDVVQAVNTSAPRMEAFSLAAVAADAPAAAGTDITNQLTNVT 240

Qy

241 VGDSGTTVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVTSKAKVPPIMAG 300

Db 241 VGDSGTTVYPHOAGYVKLYGFSVPNSAVKGDFTFKITVPKELNGLNGVTSKAKVPPIMAG 300

Qy

301 DQVLANGVIDSDGNVIYTFDYNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTT 360

XX

Db 301 DQVLAVGIDSGNVIYTFDYNTKDDYKATLTPAYIDPENVKKTGNVLTATGIGSTT 360
QY 361 ANKTVLVDYKGYKGFYNLSIKGTHIDQDKNTYRTQTIYVNSGDNVIAPIVLTGNLKPN 420
Db 361 ANKTVLVDYKGYKGFYNLSIKGTHIDQDKNTYRTQTIYVNSGDNVIAPIVLTGNLKPN 420
QY 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIPTPNQYKVEFTPD 480
Db 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIPTPNQYKVEFTPD 480
QY 481 DQITTPYIVVNGHID 496
Db 481 DQITTPYIVVNGHID 496

RESULT 12
AD89663
ID AD89663 standard; protein; 496 AA.
XX
AC AD89663;
XX
DT 20-NOV-2003 (first entry)
XX
DE Staphylococcus aureus antigenic protein #202.
XX
KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteremia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicaemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.
XX
OS Staphylococcus aureus.
XX
FN WO2003011899-A2.
XX
Hb 13-FEB-2003.
XX
FA 02-AUG-2002; 2002WO-GB003606.
XX
PR 02-AUG-2001; 2001GB-00018825.
PR 09-JAN-2002; 2002GB-00000349.
XX
PA (NYSH-) UNIV SHEFFIELD.
PA (BQS-) BIOSYNEXUS INC.
PI Foster S, Mond J, Clarke S, McDowell P, Brummel K;
PI WPI; 2003-256434/25.
DR
XX
XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
PT useful as a vaccine for immunising humans against e.g. bacteremia, septic
PT shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
PT impetigo.
XX
PS Claim 4; Page 160-161; 189pp; English.
XX
XX The present invention describes an antigenic protein or its part, which
CC is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;
CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, neuroprotective, immunosuppressive,

CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteremia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
CC bacteria-associated food poisoning, blood infections, ulcers, gonorrhoea,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, necrotising
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus antigenic
CC protein sequence from the present invention.

XX SQ Sequence 496 AA;
Query Match 53.2%; Score 2514; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.9e-120; Indels 0; Gaps 0;
Matches 496; Conservative 0; Mismatches 0;
QY 1 MNMKKEKHAIRKKSIGVASVLGTLLIGFLLSKKADASENSVTQSDSASNESKNDSS 60
Db 1 MNMKKEKHAIRKKSIGVASVLGTLLIGFLLSKKADASENSVTQSDSASNESKNDSS 60
QY 61 SVSAAPKTDDTNSDTKTSNTNNGETSVACNPAQQTTSSTSSNTATTEPTVTGEATTT 120
Db 61 SVSAAPKTDDTNSDTKTSNTNNGETSVACNPAQQTTSSTSSNTATTEPTVTGEATTT 120
QY 121 TTQANTPATQSSNTNABELVQTSNETTENDNTVSSVNSPONSNAENVSTQDTST 180
Db 121 TTQANTPATQSSNTNABELVQTSNETTENDNTVSSVNSPONSNAENVSTQDTST 180
QY 181 EATPSNNEAPQSDASNKDVVQAVNTSAPRMAFSLAAVAADAPAAAGTDTITNQLTNVT 240
Db 181 EATPSNNEAPQSDASNKDVVQAVNTSAPRMAFSLAAVAADAPAAAGTDTITNQLTNVT 240
QY 241 VGIDSGTIVVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNGLNGVTSTAKVPPIMAG 300
Db 241 VGIDSGTIVVPHQAGYVKNLNGYFVSPNSAVKGDFTFKITVPKELNGLNGVTSTAKVPPIMAG 300
QY 301 DQVLAVGIDSGNVIYTFDYNTKDDYKATLTPAYIDPENVKKTGNVLTATGIGSTT 360
Db 301 DQVLAVGIDSGNVIYTFDYNTKDDYKATLTPAYIDPENVKKTGNVLTATGIGSTT 360
QY 361 ANKTVLVDYKGYKGFYNLSIKGTHIDQDKNTYRTQTIYVNSGDNVIAPIVLTGNLKPN 420
Db 361 ANKTVLVDYKGYKGFYNLSIKGTHIDQDKNTYRTQTIYVNSGDNVIAPIVLTGNLKPN 420
QY 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIPTPNQYKVEFTPD 480
Db 421 DSNALIDQNTSIKYKVDNAADLSSEYFVNPFEDVTNSVNIPTPNQYKVEFTPD 480
QY 481 DQITTPYIVVNGHID 496
Db 481 DQITTPYIVVNGHID 496
RESULT 13
ABJ18917
ID ABJ18917 standard; protein; 913 AA.
XX
AC ABJ18917;
XX
DT 06-MAR-2003 (first entry)
XX
DE Pathogen specific antigen related staphylococcal protein SEQ ID No 63.
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW autoimmune disease; HIV; hepatitis.
XX

RESULT 15	
AAE29263	
ID AAE29263 standard; protein; 331 AA.	
XX	
AC AAE29263;	
XX	
XX	
DT 27-JAN-2003 (first entry)	
XX	
DE Staphylococcus aureus Clf33 protein.	
XX	
KW Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;	
KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;	
KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.	
XX	
OS Staphylococcus aureus.	
XX WO200272600-A2.	
PN	
XX	
PD 19-SEP-2002.	
XX	
XX 28-JAN-2002; 2002WO-US002296.	
PF	
XX 26-JAN-2001; 2001US-0264072F.	
PP	

XX	AAE29263;	
XX	AC	
XX	XX	
XX	DT	
XX	27-JAN-2003	(first entry)
XX	XX	
XX	DE	
XX	Staphylococcus aureus	Clf33 protein.
XX	XX	
XX	Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;	
XX	immunological; staphylococcal infection; impetigo; pneumonia; furuncle;	
XX	septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.	
XX	XX	
XX	Staphylococcus aureus.	
XX	OS	
XX	WO200272600-A2.	
XX	XX	
XX	19-SEP-2002.	
XX	XX	
XX	28-JAN-2002;	2002WO-US002296.
XX	XX	
XX	26-JAN-2001;	2001US-0264072P.
XX	PR	

[illegible]

PR 12-MAR-2001; 2001US-0274611P.
PR 18-JUN-2001; 2001US-0298413P.
PR 30-JUL-2001; 2001US-0308116P.
XX (INH1-) INHIBITEX INC.
PA Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
XX
XX WPI; 2002-759834/82.
DR N-PSDB; AAD46862.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
PT treating or preventing Staphylococcus aureus infection e.g. wound
PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
PT a human or animal.
XX
PS Claim 9; Page 70-72; 80pp; English.
XX
XX The invention relates to monoclonal antibody which binds the clumping
CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
CC monoclonal antibody is useful for treating or preventing S. aureus
CC infection in a human or animal, and for inhibiting the binding of
CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
CC protein is useful for inducing an immunological response in a human or
CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is Staphylococcus aureus Clf33 protein
XX
SQ Sequence 331 AA;
Query Match 36.6%; Score 1727; DB 5; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.7e-80;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 221 VAADAPAGTDTITNLTNTVTVGIDSGTTVYHQAGYVKLNYGFSVPNSAVKGDTEKITVP 280
Db 2 VAADAPAGTDTITNLTNTVTVGIDSGTTVYHQAGYVKLNYGFSVPNSAVKGDTEKITVP 61
QY 281 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKALTMPAYID 340
Db 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKALTMPAYID 121
QY 341 PENVKKTGNVTATGIGSTTANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYQTIYV 400
Db 122 PENVKKTGNVTATGIGSTTANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYQTIYV 181
QY 401 NPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIVKYKVDNADLSSEYFVNPFEDVTN 460
Db 182 NPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIVKYKVDNADLSSEYFVNPFEDVTN 241
QY 461 SVNITFPNPNQYKVFENFPDDQITTPYIVVNGHIDPNSKGDALRSTLYGYSNIIWRS 520
Db 242 SVNITFPNPNQYKVFENFPDDQITTPYIVVNGHIDPNSKGDALRSTLYGYSNIIWRS 301
QY 521 MSWDNEVAFNNGSGGDKPKVPEQPDE 550
Db 302 MSWDNEVAFNNGSGGDKPKVPEQPDE 331

Search completed: November 3, 2004, 19:02:59
Job time : 181.713 secs

QY	33	SKKHA	DAS	ENV	TQ	SD	AS	NE	---	SK	ND	SS	V	S	A	P	K	T	D	T	N	V	S	D	T	K	T	S	S	T	N	N	G	E	T	S	V	89			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	756	SOQ	F	P	F	F	A	N	I	L	D	N	O	P	A	S	Q	T	M	S	P	A	H	V	S	S	P	V	P	A	Q	M	O	Q	O	H	T	Q	O	P	815
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
QY	90	AQ	N	P	A	Q	E	I	T	S	S	T	---	N	A	T	T	E	T	P	T	G	E	A	T	T	T	T	T	T	T	T	T	T	T	T	T	144			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	816	SQ	-	P	O	Q	T	O	Q	O	P	S	T	P	S	H	L	S	I	L	N	E	T	A	N	S	D	S	T	L	A	V	---	A	N	D	S	871			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
QY	145	TG	N	E	T	T	F	N	D	T	V	S	S	V	N	S	P	O	N	S	T	A	E	N	V	S	T	Q	D	T	S	T	E	A	T	S	203				
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	872	-	N	O	P	T	S	V	H	D	S	O	S	S	I	N	S	S	N	---	N	O	Q	L	L	M	D	N	T	S	H	G	G	S	E	S	925				
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
QY	204	Q	A	N	T	S	A	P	R	M	R	A	P	S	L	A	A	A	D	A	P	A	G	T	D	T	N	O	L	T	---	V	T	V	I	D	249				
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	926	MS	N	S	A	A	---	V	A	A	S	N	A	A	A	A	L	D	N	S	O	S	S	N	A	E	F	F	K	N	O	S	G	E	G	L	G	981			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
QY	250	Y	P	H	Q	A	G	Y	V	K	L	N	Y	G	F	S	V	---	P	N	S	A	V	K	---	G	T	E	K	---	I	T	V	P	K	E	L	285			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	982	L	P	O	D	E	N	S	V	K	T	O	I	T	I	S	T	E	L	P	T	O	D	L	E	K	P	T	A	E	M	E	Q	T	O	L	G	1041			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
QY	286	N	G	V	T	S	T	A	K	V	---	P	I	M	A	G	D	O	V	L	A	N	---	V	I	D	S	---	---	---	---	---	---	---	---	---	---	---	311		
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	1042	S	G																																						

[illegible]

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 45.7559 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2
Perfect score: 4725
Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSLFLFRKKENKDKK 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4725	100.0	933	3	US-08-293-728-2
2	4725	100.0	933	3	US-09-421-868-2
3	4662	98.7	936	4	US-08-956-171E-5249
4	4662	98.7	936	4	US-08-781-986A-5249
5	1810.5	38.3	918	4	US-08-200-850E-1
6	1727	36.6	345	3	US-08-856-253-7
7	1628	34.5	1092	4	US-09-147-405B-15
8	1584.5	33.5	1742	4	US-09-386-962C-4
9	1584.5	33.5	1742	4	US-09-386-959-4
10	1414	29.9	1166	4	US-09-200-650E-7
11	1241.5	26.3	930	4	US-09-200-650E-3
12	1110.5	23.5	1315	4	US-09-200-650E-5
13	880	18.6	930	3	US-09-134-001C-5314
14	880	18.6	930	4	US-09-386-962C-10
15	847.5	17.9	2137	3	US-09-134-001C-4463
16	799	16.9	287	4	US-09-710-279-468
17	695	14.7	486	4	US-09-710-279-788
18	671	14.2	487	4	US-09-386-962C-14
19	671	14.2	487	4	US-09-386-959-65
20	614	13.0	1027	4	US-08-956-171E-5254
21	614	13.0	1027	4	US-08-781-986A-5254
22	530.5	11.2	1155	4	US-09-710-279-1780
23	488	10.3	466	3	US-09-134-001C-4749
24	472.5	10.0	669	4	US-09-107-532A-6532
25	452	9.6	582	4	US-09-147-405B-13
26	452	9.6	593	4	US-09-147-405B-11
27	413	8.7	414	4	US-09-248-796A-19046

28	388.5	8.2	556	4	US-09-248-796A-22338	Sequence 22338, A
29	379	8.0	2870	4	US-09-479-467A-15	Sequence 15, Appl
30	379	8.0	3178	4	US-09-479-467A-4	Sequence 4, Appl
31	350	7.4	461	4	US-09-248-796A-23039	Sequence 23039, A
32	346	7.3	827	4	US-09-248-796A-17307	Sequence 17307, A
33	344.5	7.3	1060	3	US-08-911-393-2	Sequence 2, Appl
34	344.5	7.3	1060	4	US-09-955-909-2	Sequence 44, Appl
35	343.5	7.3	447	4	US-09-794-422-44	Sequence 2, Appl
36	334.5	7.1	886	4	US-08-956-171E-5235	Sequence 44, Appl
37	334.5	7.1	886	4	US-08-781-986A-5235	Sequence 5235, Ap
38	319.5	6.8	699	4	US-09-248-796A-23200	Sequence 5235, Ap
39	302.5	6.4	917	4	US-09-248-796A-18960	Sequence 23200, A
40	300	6.3	1776	4	US-09-556-877-179	Sequence 18960, A
41	300	6.3	1776	4	US-09-620-412C-179	Sequence 179, App
42	300	6.3	1776	4	US-09-598-419-179	Sequence 179, App
43	299.5	6.3	821	4	US-09-556-877-195	Sequence 195, App
44	299.5	6.3	821	4	US-09-620-412C-195	Sequence 195, App
45	299.5	6.3	821	4	US-09-598-419-195	Sequence 195, App

ALIGNMENTS

RESULT 1
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match	100.0%;	Score	4725;	DB	3;	Length	933;
Best Local Similarity	100.0%;	Pred. No.	1.9e-293;				
Matches	933;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
Qy	1	MNMKKEKHAIRKKSIGVASVLVGLTGLGFLLSKEADASENSVTQSDSASNESKNDSS	60				
Db	1	MNMKKEKHAIRKKSIGVASVLVGLTGLGFLLSKEADASENSVTQSDSASNESKNDSS	60				
Qy	61	SVGAAPKTDITNYSDDTKTSNTNNGETSVACNPAQOETTSSTTNAETPTVTGEATTT	120				
Db	61	SVGAAPKTDITNYSDDTKTSNTNNGETSVACNPAQOETTSSTTNAETPTVTGEATTT	120				
Qy	121	TTQANTPATTSQSSNTNAELVNQTSNETTFNDNTVSSVNSPONSNAENVTQDTST	180				
Db	121	TTQANTPATTSQSSNTNAELVNQTSNETTFNDNTVSSVNSPONSNAENVTQDTST	180				
Qy	181	EATPSNNESAPQSDASNKDVNQAVNTSAPRMRFAFSLAAVAADAPAAAGDITNQLTNVT	240				
Db	181	EATPSNNESAPQSDASNKDVNQAVNTSAPRMRFAFSLAAVAADAPAAAGDITNQLTNVT	240				
Qy	241	VGIDSGTTVPHQAGYVKNLNGYFVSPNSAVKGTFTKTPKELNNGVTSTAKVPPIMAG	300				
Db	241	VGIDSGTTVPHQAGYVKNLNGYFVSPNSAVKGTFTKTPKELNNGVTSTAKVPPIMAG	300				
Qy	301	DQVLANGVIDSGNVYITFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTT	360				
Db	301	DQVLANGVIDSGNVYITFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTT	360				
Qy	361	ANKTVLVDYKYGKFNLSIKGTIDQDKTNNTYRQIYVNPSPGDNVIAPIVLTGNLKPT	420				
Db	361	ANKTVLVDYKYGKFNLSIKGTIDQDKTNNTYRQIYVNPSPGDNVIAPIVLTGNLKPT	420				

QY 421 DSNALIDQONTISIKVYKVDNAADLSESYFVNPENFEDVTNSVNIITPPNPQYKVFETPD 480
DB 421 DSNALIDQONTISIKVYKVDNAADLSESYFVNPENFEDVTNSVNIITPPNPQYKVFETPD 480
QY 481 DQITTPYIIVVNGHIDPNSKGLALRSTLYGYNSNIIWRSMWDNEVAFNNGSGSDGID 540
DB 481 DQITTPYIIVVNGHIDPNSKGLALRSTLYGYNSNIIWRSMWDNEVAFNNGSGSDGID 540
QY 541 KPVVPEQDEPEGEIPIPEDSDPGSDGSDNSDSDSDSDSDSDSDSDSDSDSDSDSD 600
DB 541 KPVVPEQDEPEGEIPIPEDSDPGSDGSDNSDSDSDSDSDSDSDSDSDSDSDSDSD 600
QY 601 SDSDSASD 660
DB 601 SDSDSASD 660
QY 661 SD 720
DB 661 SD 720
QY 721 SD 780
DB 721 SD 780
QY 781 SD 840
DB 781 SD 840
QY 841 SD 900
DB 841 SD 900
QY 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933
DB 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933

RESULT 2
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2
Query Match 100.0%; Score 4725; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.9e-293;
Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNMCKKHAIRKKSIGVASVLVGLTIGFLLSKEADASENSVTQSDASNEKSDSS 60
DB 1 MNMCKKHAIRKKSIGVASVLVGLTIGFLLSKEADASENSVTQSDASNEKSDSS 60
QY 61 SVSAAKPTDDTNVSTKTSNTNNGETSVANPAQQTQSSNTATTEPTVGEATT 120
DB 61 SVSAAKPTDDTNVSTKTSNTNNGETSVANPAQQTQSSNTATTEPTVGEATT 120
QY 121 TTNQANTPATTQSSNTNAEELVNOTSETTENDTNTVSSVNSPQNSTNAENVSTQDTST 180

DB 121 TTNQANTPATTQSSNTNAEELVNOTSETTENDTNTVSSVNSPQNSTNAENVSTQDTST 180
QY 181 EATPSNNEAPOSTDASHKDVQVNOAVNTSAPRMRAFSLAAVAADAPAACTDTITNOLTNVT 240
DB 181 EATPSNNEAPOSTDASHKDVQVNOAVNTSAPRMRAFSLAAVAADAPAACTDTITNOLTNVT 240
QY 241 VGIDSGTTVPHQAGVYKLVNGFSPNAGVKGDTFKITVPKELNNGVTSTAKVPPIMAG 300
DB 241 VGIDSGTTVPHQAGVYKLVNGFSPNAGVKGDTFKITVPKELNNGVTSTAKVPPIMAG 300
QY 301 DOVLANGVIDSDGNVIYTFDTYVNTKDDVKATLTMPAYIDPENVKKTGNVTIATGISTP 360
DB 301 DOVLANGVIDSDGNVIYTFDTYVNTKDDVKATLTMPAYIDPENVKKTGNVTIATGISTP 360
QY 361 ANKTVLVYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVITAPVLITGLMKPT 420
DB 361 ANKTVLVYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVITAPVLITGLMKPT 420
QY 421 DSNALIDQONTISIKVYKVDNAADLSESYFVNPENFEDVTNSVNIITPPNPQYKVFETPD 480
DB 421 DSNALIDQONTISIKVYKVDNAADLSESYFVNPENFEDVTNSVNIITPPNPQYKVFETPD 480
QY 481 DQITTPYIIVVNGHIDPNSKGLALRSTLYGYNSNIIWRSMWDNEVAFNNGSGSDGID 540
DB 481 DQITTPYIIVVNGHIDPNSKGLALRSTLYGYNSNIIWRSMWDNEVAFNNGSGSDGID 540
QY 541 KPVVPEQDEPEGEIPIPEDSDPGSDGSDNSDSDSDSDSDSDSDSDSDSDSDSDSD 600
DB 541 KPVVPEQDEPEGEIPIPEDSDPGSDGSDNSDSDSDSDSDSDSDSDSDSDSDSDSD 600
QY 601 SDSDSASD 660
DB 601 SDSDSASD 660
QY 661 SD 720
DB 661 SD 720
QY 721 SD 780
DB 721 SD 780
QY 781 SD 840
DB 781 SD 840
QY 841 SD 900
DB 841 SD 900
QY 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933
DB 901 SEDEANTSLIWGLLASIGSLLLFRKKENKDKK 933

RESULT 3
US-08-956-171E-5249
; Sequence 5249, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Patrick S. Dillon
; Gil H. Choi
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland

COUNTRY: USA
ZIP: 20850
COMPUTER/READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5249:

Query Match 98.7%; Score 4662; DB 4; Length 936;
Best Local Similarity 99.0%; Pred. No. 2e-289;
Matches 924; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY	1	MNNKKKEKHAIKKSIGVASVLGTLIGFLLGSKKEADASENSVTQSDASNESKNDSS	60
Db	10	MNNKKKEKHAIKKSIGVASVLGTLIGFLLGSKKEADASENSVTQSDASNESKNDSS	69
QY	61	SVSAAPKTDVTSNNTNNGETSVQNPQAQQTSSSTNATTEPTVTGEATT	120
Db	70	SVSAAPKTDVTSNNTNNGETSVQNPQAQQTSSSTNATTEPTVTGEATT	129
QY	121	TNQAANTPATQSSNTNABELVNGTSTNNTNNGETSVQNPQAQQTSSSTNATTEPTVTGEATT	180
Db	130	TNQAANTPATQSSNTNABELVNGTSTNNTNNGETSVQNPQAQQTSSSTNATTEPTVTGEATT	189
QY	181	EATPSNNEAPOSTDASNDKVQAVNTSAPRPAFSLAAVAADAPAAAGTDITNQLNVT	240
Db	190	EATPSNNEAPOSTDASNDKVQAVNTSAPRPAFSLAAVAADAPAAAGTDITNQLNVT	249
QY	241	VGIDSGTTPVPHQAGYVKLYGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPPIMAG	300
Db	250	VGIDSGTTPVPHQAGYVKLYGFSVPNSAVKGDFTKITVPKELNGLNGVTSTAKVPPIMAG	309
QY	301	QVILANGVIDSDGNVIITFTDYVNTKDDVKAITLTPAYIDPENVKKTGNVLTATIGSTT	360
Db	310	QVILANGVIDSDGNVIITFTDYVNTKDDVKAITLTPAYIDPENVKKTGNVLTATIGSTT	369
QY	361	ANKTLVLDYKYGKFNLSIKGTIDQDKTNTYRQTIYVNPNGDNVIAVLTGNLKPNT	420
Db	370	ANKTLVLDYKYGKFNLSIKGTIDQDKTNTYRQTIYVNPNGDNVIAVLTGNLKPNT	429
QY	421	DSNALIDQQNTSIKVKYVDNAADLSESYFVNPNFEDVTNSVNTTFPNPQYKVFNTPD	480
Db	430	DSNALIDQQNTSIKVKYVDNAADLSESYFVNPNFEDVTNSVNTTFPNPQYKVFNTPD	489
QY	481	DQITTPYVVVNGHIDPNSKGLALRSTLYGYNSTIIWRNSWSDNEVAFNNGSGSGDID	540
Db	490	DQITTPYVVVNGHIDPNSKGLALRSTLYGYNSTIIWRNSWSDNEVAFNNGSGSGDID	549

QY	541	KPVVPEQDPDPGEIEPIPEDSDSDPGSDSGSDNSGSDSGSDSTSDSGSDSASDSDSA	600
Db	550	KPVVPEQDPDPGEIEPIPEDSDSDPGSDSGSDNSGSDSGSDSTSDSGSDSASDSDSA	609
QY	601	SDSDSASDSDSASDSDSASDSDSDNDSDSDSDSDSDSDSDSDSDSDSDSDSDSD	660
Db	610	SDSDSASDSDSASDSDSASDSDSDNDSDSDSDSDSDSDSDSDSDSDSDSDSDSD	669
QY	661	SD	720
Db	670	SD	729
QY	721	SD	780
Db	730	SD	789
QY	781	SD	840
Db	790	SD	843
QY	841	SDSGSDSDSDSDSDSDSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSG	900
Db	844	SDSGSDSDSDSDSDSDSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSG	903
QY	901	SEDEANTSLIWGLLASIGSILLFFRRKKENKDKX	933
Db	904	SEDEANTSLIWGLLASIGSILLFFRRKKENKDKX	936

RESULT 4

US-08-781-986A-5249
Sequence 5249, Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:

CLASSIFICATION:

435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 5249:

SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match

98.7%; Score 4662; DB 4; Length 936;

Best Local Similarity 99.0%; Pred. No. 2e-289;
Matches 924; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 1 MNMKKEKHAIRKKSIGVASVLVGLTIGFGLSSKEADASENSVTQSDASNESKSDSS 60
DB 10 MNMKKEKHAIRKKSIGVASVLVGLTIGFGLSSKEADASENSVTQSDASNESKSDSS 69
QY 61 SVSAAPKTDITNVSDTKSSNTNNGETSVAQNPQAQETTSQSSNTATTEETPVTEATT 120
DB 70 SVSAAPKTDITNVSDTKSSNTNNGETSVAQNPQAQETTSQSSNTATTEETPVTEATT 129
QY 121 TTNQANTPATTOSSNTNABELVQTSNETTFNDTNTVSSVNSPQNSNAENVSTTQDST 180
DB 130 TTNQANTPATTOSSNTNABELVQTSNETTFNDTNTVSSVNSPQNSNAENVSTTQDST 189
QY 181 EATPSNNEAPOSTDASNDKVVNQAVNTSAPRMRAFSLAAVAADAPAGTDTITNLTNVT 240
DB 190 EATPSNNEAPOSTDASNDKVVNQAVNTSAPRMRAFSLAAVAADAPAGTDTITNLTNVT 249
QY 241 VGIDSGTTVYVPHQAGVYKLVNFGVSPNSAVKGDTEFKITVPKELNLNGVTSIAKVPPIMAG 300
DB 250 VGIDSGTTVYVPHQAGVYKLVNFGVSPNSAVKGDTEFKITVPKELNLNGVTSIAKVPPIMAG 309
QY 301 DQVLANGVIDSGNVIIYFTDYVNTKDDVKAATLTPAYIDPENVKKTGNVTLATGIGSTT 360
DB 310 DQVLANGVIDSGNVIIYFTDYVNTKDDVKAATLTPAYIDPENVKKTGNVTLATGIGSTT 369
QY 361 ANKTVLVDEYKGVKFNLSIGTIDQIDKTNNTVEQTIYVPSGDNVIAVLITGNLKENT 420
DB 370 ANKTVLVDEYKGVKFNLSIGTIDQIDKTNNTVEQTIYVPSGDNVIAVLITGNLKENT 429
QY 421 DSNALIDQNTSIKVKYVDNAADLSESYFVNPFEDVTNSVNIITFPNPQYKVFNTPD 480
DB 430 DSNALIDQNTSIKVKYVDNAADLSESYFVNPFEDVTNSVNIITFPNPQYKVFNTPD 489
QY 481 DOIITPYIVVNGHIDPNSKGLALRSTLYGYNINIIWRMSWMDNEVAFNNGSGGDGID 540
DB 490 DOIITPYIVVNGHIDPNSKGLALRSTLYGYNINIIWRMSWMDNEVAFNNGSGGDGID 549
QY 541 KPVPPEQDEPEGEIEPIEDSDSDPGSDSGSDSDSDSDSDSDSDSDSDSDSDSDSD 600
DB 550 KPVPPEQDEPEGEIEPIEDSDSDPGSDSGSDSDSDSDSDSDSDSDSDSDSDSDSD 609
QY 601 SDSASD 660
DB 610 SDSASD 669
QY 661 SDSASD 720
DB 670 SDSASD 729
QY 721 SDSASD 780
DB 730 SDSASD 789
QY 781 SDSASD 840
DB 790 SDSASD 843
QY 841 SDSASD 900
DB 844 SDSASD 903
QY 901 SEDEANTSLIWGLLASIGSLIIIFRKKNKDKK 933
DB 904 SEDEANTSLIWGLLASIGSLIIIFRKKNKDKK 936

RESULT 5

US-09-200-650E-1

; Sequence 1, Application US/09200650E

; Patent No. 6680195

; GENERAL INFORMATION:

; APPLICANT: Patti, Joseph M.
; APPLICANT: Poster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eidhinn, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 918
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-200-650E-1

Query Match 38.3%; Score 1810.5; DB 4; Length 918;
Best Local Similarity 43.5%; Pred. No. 1.8e-107;
Matches 429; Conservative 148; Mismatches 268; Indels 141; Gaps 31;

QY 5 KKEKHAIRKKSIGVASVLVGLTIGFGLSSKEADASE--NSVTQSDASNESKSDSSV 62
DB 16 KQKVISIRFTVGTTSVIVGATILFG-IGNHQAQASEQSDNTTQS-SKNNASADEKNNM 73
QY 63 SAAPKTDITNVSDTKSSNTNNGETSVAQNPQAQETTSQSSNT-ATTETPVTEATT 121
DB 74 IETPQNTTANDTSDISANTNSANVDSTTKEMSTQTSNTTTEPASTNETP---QPTAI 129
QY 122 TTNQANTPATTOSSNTNABELVQTSNETTFNDTNTVSSVNSPQNSNAENVSTTQDST 181
DB 130 KNOAT--AAKMQDQTVQEGNSQVDNKT-NDANSI-----ATNSE---LNKSQL 174
QY 182 ATPSNNEAPOSTDASNDKVVNQAVNTSAPRMRAFSLAAVAADAP-----AAGTDITNQ 235
DB 175 DLP---QSSPT-----ISNAQGTSKPSVRFVAVRSLAVAEFVVAADAKGTVNDK 223
QY 236 LTNTVGTIDSGTTVYVPHQAGVYKLVNFGVSPNSAVKGTPTKITVPKELNLNGVTSIAKVP 295
DB 224 VTASNFKLEK-TTFDPNQSGNTFMAANFTVTDKVKSGDYF-----TAKLP 267
QY 296 PIMAGDQVLANGVID-----SDGNVI-----YTFDYVNTKD 327
DB 268 ----DSLGTGSDVDYSSNNMTPIADIKSTNGDVAKATYDILTKTVTFVFTDYNNKE 322
QY 328 DVKATLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDEYK---YGFYNLISKGT 383
DB 323 NINGQFSLPFTDRAPKAPKSGTYDANINIADEMFMFNKITNYSSPIAGIDKPGNANISQ 382
QY 384 IDQIDKNT--NTYQTIYVNSP-----GDNVIAVLITGNLKENTDSNALIDQNTSIKVKY 437
DB 383 IIGVDTASGQNTYQTVFVNPQKRVLGNTWV--YIKGYODKIERSSGKVSATDKLRIFE 440
QY 438 VDNAADLSESYFVNP--ENPEDVNSV--NITFPNPQYKVFNTPDQITPIYVNVNG 493
DB 441 VNDTSKLSDSYVADPNSDNLKEVTDQFKRIYYEHPNVASIKFG---DITKTYVVLVEG 496
QY 494 HIDPNSKGLALRSTLYGYN-----SNIWRMSWMDNEVAFNNGSGSDGIDKPVVPEOPD 549
DB 497 HYDNTGKN---LKTQVIQENVDPVNTDRDYSIFGWNENNVVYGGGADG--DSAVNPKDPT 552
QY 550 EPGIEPIPE--DSDDPGSDSGSDNSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 608
DB 553 PGPPVDPPEPDPPEPTD-----PEPSPDPEPSPDPPDSDSD 594
QY 609 SDSASD 668
DB 595 SDSASD 654

[illegible]

RESULT 6
 US-08-856-253-7
 ; Sequence 7, Application US/08856253
 ; Patent No. 6288214
 ; GENERAL INFORMATION:
 ; APPLICANT: Hook, Magnus
 ; APPLICANT: Patti, Joseph M.
 ; APPLICANT: House-Pompeo, Karen
 ; APPLICANT: Sthanam, Narayana
 ; APPLICANT: Symersky, Jindrich
 ; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
 ; TITLE OF INVENTION: AND METHODS OF USE
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: U.S.
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/856,253
 ; FILING DATE: Concurrently Herewith
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/017,678
 ; FILING DATE: 16-MAY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: TANK:193
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (512) 474-7577
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 345 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 US-08-856-253-7

Db	14	VAADAPAAAGTDTINQLNNTVTGIDSGTTVYPHQAGYVKLNYGFSVPNSAVKGDTFKKTIVP	73
Qy	281	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	340
Db	74	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	133
Qy	341	PENVKKTGNVTLTATIGSTTTANKTTLVDYBKYGKFFYNLSIKGTIDQIDKTNNTVRQTIYV	400
Db	134	PENVKKTGNVTLTATIGSTTTANKTTLVDYBKYGKFFYNLSIKGTIDQIDKTNNTVRQTIYV	193
Qy	401	NPSGDNVIAVPLTGNLKPNTDSNALIDQONTSTKVYKVDNAADLSESYFVNPENFEVDTN	460
Db	194	NPSGDNVIAVPLTGNLKPNTDSNALIDQONTSTKVYKVDNAADLSESYFVNPENFEVDTN	253
Qy	461	SVNITTFPNQYKVFENTPDDQITTPPIYVVVNGHIDPNSKGDALRSTLYGYSNIIWRS	520
Db	254	SVNITTFPNQYKVFENTPDDQITTPPIYVVVNGHIDPNSKGDALRSTLYGYSNIIWRS	313
Qy	521	MSMDNEVAFNNGSGSGDIDKPVVPEQDPE	550
Db	314	MSMDNEVAFNNGSGSGDIDKPVVPEQDPE	343

RESULT 7
US-09-147-405B-15
Sequence 15, Application US/09147405B
Patent No. 6733758
GENERAL INFORMATION:
APPLICANT: Guss, Bengt
APPLICANT: Nilsson, Martin
APPLICANT: Frykberg, Lars
APPLICANT: Flock, Jan-Ingmar
TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
Coagulase-Negative Staphylococcus
FILE REFERENCE: guss 09/147405
CURRENT APPLICATION NUMBER: US/09/147,405B
CURRENT FILING DATE: 1999-04-11
PRIOR APPLICATION NUMBER: PCT/SE97/10191
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: SE 9602496-3
PRIOR FILING DATE: 1996-06-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 1092
TYPE: prt
ORGANISM: Staphylococcus epidermidis
US-09-147-405B-15

Query Match	36.6%	Score 1727;	DB 3;	Length 345;
Best Local Similarity	100.0%;	Pred. No. 1.1e-102;		
Matches 330;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	221	VAADAPAAAGTDITNQLNVTNVGIDSGTTPVPHOAGYVKLYGFSVPNSAVKGDTEFKITVP	280	
Qy	192	QSTDASNKDVVNQAVNTSAPRMFAFSLAAVAADAAPAGTDTINQLTNVTVGI-DSGTIVY	250	
Db	197	KESTESGKEENTIEQPNVKEDSTTSQPSGYNTDEKISNQDELLN-LPINEYENKARP	255	
Qy	144	QISNEIEFNDIINIV-----SSVNSQFNSINENSVISIQDLSIERAFPN-----NCARF	179	

```
Db 256 LSTTSAQPSIKRVTWN-----QLAAEQSGSNVHLLIKVTDQ--SITEGYDDSEGVK 304
QY 251 PHOAGVVKLYNGFSPNSAVKGDTFKIVPKELNNGVTSTAKVPIIMAGD-OVLANGVI 309
Db 305 ABDAENLIYDVTFEVDKVKSGDGTMTVDIDKNTVPSDLTSTPIKIDNSGEIATGY 364
QY 310 DS-DGNVIYFTDYVNTKDDVKATILTPAYIDPENVKKTG---NVTLATGIGSTTANKTV 365
Db 365 DNKNQIITVFTDYVDYKNIKAHLKLTSDYDKSVNNNTKLDVEYKTALSS--VNKTI 422
QY 366 LVDYKFKYFNLSKIGTIDQIDKNNVTRQIYVNPNGDNVIAPIVLTGNLKPNTSNAL 425
Db 423 TVEYQRENRTANLQSMFTNIDTKNHTVEQIYINPL--RYSAKETVNIISGNGDEGT 480
QY 426 IDQWNTSLKVKYVDAADLSEYFV-NPENFEDVNSVNIITPNNQYKVENTPDQIT 484
Db 481 IIDDSTIIKVKYVGNQNLPSNRIDYSEYEDVNDYVAQIGNNNVNNINFG---NID 536
QY 485 TPYIYVNGHIDPNSKGD-----LALRSTLYGYSNIIIRMSGMDNEVAFNNGSGSD 537
Db 537 SPYIIKVISKYDPN-KDDVTYTIQQTVMQTTINEYTG--PRTASVDNTIAFTSSGQOQ 593
QY 538 G-----IDKVPV-----EOP-----DEPGEIE 555
Db 594 GDLPPKTYKIGYVWEDVDGQIQNTNDNEKPLSNVLVLTYPDGTSKSVRTDEGKYQ 653
QY 556 -----PIPEDSDSDPGSDSGSDS-----NSDSDSG- 582
Db 654 FDCLKNGLTYKITFTPEGYFTLKHSGTNPALDSEGSVWVTINGQDDMTIDSGFYQTP 713
QY 583 ----- 582
Db 714 KYSLGNVYVYDNTKDGIOGDDKEKIGVKVTLKDENGNIISTTTTDENGKYQFDNLNSGN 773
QY 583 -----SDSTSDSGSDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 631
Db 774 YIVHFKPSGMTQTITDSDGDDDEQADGEEVHTTIDHDDFIDNGYYDDSDSDSDSDS 833
QY 632 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 691
Db 834 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 893
QY 692 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 751
Db 894 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 953
QY 752 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 811
Db 954 DSDSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1009
QY 812 DSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 871
Db 1010 DSVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1046
QY 872 VPPNPKNGTNASNNKNEAKSKEPLPDGTGSEDE--ANTSLIWGLLASTGSLILFRKXEN 929
Db 1047 -----KSTKDKLPDGTGANEDYSGKGLTLGLTFLAGLALLGKRRKRN 1088
QY 930 KDK 932
Db 1089 KKK 1091
```

RESULT 8

```
US-09-386-962C-4
; Sequence 4, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: POSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
```

```
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1742
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-386-962C-4

Query Match 33.5%; Score 1584.5; DB 4; Length 1742;
Best Local Similarity 30.6%; Pred. No. 1e-92;
Matches 452; Conservative 156; Mismatches 285; Indels 583; Gaps 35;

QY 5 KKEKHAIRKKSIGVASVLVGLTIGLLSSKKAADASNSVTSQSDASNE-----SKNDSBS 60
Db 28 KVKYSIRKFTVGTASILVGLATMFG-AADNEAKAAEDN--QLESASKEEQKSGRDNENS 84
QY 61 SVS-----AAPKTD-----TNVS-DTTSNT 82
Db 85 KLNQVDLNGSHSSEKTNVNNATEVKVEATTSVSKPKANEAVVNESKPKTTEAP 144
QY 83 NNGETSVAQN-----AQOETTQSS-----STNATTEETPTGEATT-----TTTNOAN 126
Db 145 TVNEESIAETPKSTTTQDSTIEKNPDLKDLNLSSTTSKESKTDHSTKQACMSTKN 204
QY 127 ---TPATTCSSNTNABELVNQTSNE-----TTFND----- 153
Db 205 LDTNDSPTQSEKTSQANNSTDNQSAFSLQDSDKPSQKVVYKTFNDEFTQDVEHTTK 264
QY 154 -----TNTVSSVNSPQNSTNAENVSTTQDTS-----TEATPSNESAPQSTDAASKOVVN 203
Db 265 LKTPSVSTDSVNDKQDYTRSAVASLGVDSNETEAITNAVRDNLDLKAASRQIIEAIIA 324
QY 204 QAV-----NTSAPRMAFSLAAVAADAPAAAGTDITN-- 234
Db 325 EALKKDFSNPDYGVDTPLALNRSQSKNSPHKSASPRMLMSLAA---EPNSGKNVNDKV 380
QY 235 QLTNVTVGID-----SGTIVPHQAGYVKLYNGFVSPNSAVKGDTFKIVPKELNNGVT 289
Db 381 KITNPTLSLNSNNHANNVIMFTSNEQFNKANYELDSDSIKEGDTFTIKYGYIRPGGLE 440
QY 290 STAKVPIIMAGD-QVLANGVIDSDGN-VIYFTDVTNKKDVKATLTWPAIDPN-VKK 346
Db 441 LPAIKTQURSKDGSIVANGVYDKTNTTNTTYFTNTVDQYQNTITGSFDLIATPKRTAKD 500
QY 347 TGNVTLATGIGSTTANKTVLDYKFKYFNLSIKGTIDQIDKNTTYRQIITYVNPNGDN 406
Db 501 NQNYEMEVTIANEVVKDFIVD---YGNKKDNTTTAAVANVDVNNKNEVVYLLNQNN 557
QY 407 -----VIAPVLITGNLKN-----TDSNALIDQNTSI----- 433
Db 558 PKYAKYFTSVKNGEPIGGEVKYVEYVTDNAMVDSFNPDLNSNVKDVTSQFAPKVSADGT 617
QY 434 -----KVYKVDNAAD-----LSSEYFVNPNFEDVNSVNIITPN 468
Db 618 RVDINFARSMANGKXIIVTQAVRPTGCTGNVTEYWLTRDGTNTNDFRGYKSTIVTYLN 677
QY 469 PNQ----- 471
Db 678 GSSTAQGDNPYSLGDYVWLDKKNKGVDDEKGLAGVYVTLKDSNNRELQRTVTTDQSGH 737
QY 472 -----YKVEENTPDDQITTP-----YIVVNGHI----- 495
Db 738 YQFNLQNGTYTVEFAIPDNTYTPSPANNSTDAIDSDGERDGTGRKVVYAKGTINNADNMT 797
QY 496 -----DPNSKGDIALRSTLYGYNSNIIRMSGMDNE----- 526
Db 798 VDTGYLFTKYNVGVYVWEDNKGQIQDNEKGISGVKVTLLKKNKGDITIGTITDNGKY 857
QY 527 --VAFNNG-----SGSGDGDID-----KP- 542
```


Db 858 EFTGLENGDYTIEFETPGYPTTKQNSGDSBGKDSNGTKITVTYVKDADNKTIDSGFYKPT 917
QY 543 -----
Db 918 YNLGDYVWEDTNKDGIDQDSEKIGISGVKVTLLKDKNGNAIGTTTTTASGHYQFKGLENGSY 977
QY 543 -VVEQP-----
Db 978 TVEFETPGYPTKANSQDITVDSNGITTTGIINGADNLTIDSGFYKTPKYSVGDYVWE 1037
QY 549 -----DEPGEIEPI-----
Db 1038 DTNKGIDQDNEKIGISGVKVTLLKDKNGNAIGTTTTTASGHYQFKGLENGSY 1097
QY 560 -----
Db 1098 MTQTTANSNGNDEKADGEDVRVTITDHDDFSIDNGYFDDSDSDSDSDSDSDSDSDSDA 1157
QY 572 DSNDSGSD 631
Db 1158 DSD 1217
QY 632 DSD 691
Db 1218 DSD 1277
QY 692 DSD 751
Db 1278 DSD 1337
QY 752 DSD 811
Db 1338 DSD 1397
QY 812 DSD 871
Db 1398 DSD 1457
QY 872 VPENSPKNGTNASNKAKEPLPTGSEDEANT 907
Db 1458 DSD 1493

RESULT 9
US-09-386-959-4
; Sequence 4, Application US/09386959
; Patent No. 6703025
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; APPLICANT: FOSTER, Timothy J.
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: MULTICOMPONENT VACCINES
; FILE REFERENCE: P06333USI/BAS
; CURRENT APPLICATION NUMBER: US/09/386,959
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,439
; EARLIER FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1742
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-386-959-4

Query Match 33.5%; Score 1584.5; DB 4; Length 1742;
Best Local Similarity 30.6%; Pred. No. 1e-92;
Matches 452; Conservative 156; Mismatches 285; Indels 583; Gaps 35;
QY 5 KKEKHAIRKXISGVASVILGTIGFLLSKSEADASENVTSQDSASNE-----SKSNDSS 60
Db 28 KVKYSIRKFTVGTASILVGTATLMFG-AADNEAKAEDN--QLESASKEQKGRDNE 84

QY 61 SVS-----AAPKTD-----TNVS-DTKTSST 82
Db 85 KLNQVDLDNGSHSEKTTNNNAATEVKKVEAPTISDVSKPANEAIVTNESKTKTEAP 144
QY 83 NNGETSAQNP-----AQETTQS-----STNATTEETPTVTEATT-----TTNQAN 126
Db 145 TVNEESIAETPKTSTTQODSTERNPSLKDNLSSSTTSKESKTDHSTKQAOASTKSN 204
QY 127 ---TPATTQSNTNAELVNQTSNE-----TTFND----- 153
Db 205 LDTNDSPTQSEKTSQANNDSTNQAPSQLOSKPSEQVYKTKFNDPTQDVEHTTK 264
QY 154 ---TNVSNSPONSNAENVSTTQDTS-----TEATPSNESAPOSTDASNDKDVN 203
Db 265 LKTPSVSDSDSDKQDYTRSAVASLGVDSNETAITNAVRDNLJLKAAREQINEAIIA 324
QY 204 QAV-----NTSAPRMRAFSLAAVAADAAPAGTDITN-- 234
Db 325 EALKKDFSNPDYGVDTPLALNRQSCKNSPHKSASPRNLMSLAA---EPNSGKNVNDKV 380
QY 235 QLTNVTGID-----SGTIVYHQAGYVKLYGSPVNSAVKGTFTKITVPKELNNGVT 289
Db 381 KITNPTLSLKNHANNVWP*SNQNLKANYELDDSIKEGDTFTIKYQYIRPGGLE 440
QY 290 STAKVPPIMAGD-QVLANGVIDSDGN-VIYTFDYNTKDDVKATLTMPAVIDPEN-VKK 346
Db 441 LPAIKTLQRLSKDGSIVANGYVDKTTNTTTFYFNVDYQYQITGSDFLIATPKRETAIKD 500
QY 347 TGNVTLATIGSTTANKTLVLVDEYKGYKFNLSIKGTIDQIDKTNNTYRTIYVNSGDN 406
Db 501 NQNVEMVTIANEVKKDFIVD---YGNKDNITTAANVNDVNNKHNEVYLNQNNQ 557
QY 407 ---VIAPVLTGMLKN-----TDSNALIDQNTSI----- 433
Db 558 PKYAKYFTSVKNGEFIPGEVKVVEVDTNAMVDSFNDLNSNVKDVTSQFAPKVSADGT 617
QY 434 ---KVKVVDNAAD-----LSESYFVNPENEDVNSNITFEN 468
Db 618 RVDINFARSANGKKYIVTQAVRPTGNGVYTEWLTFRDGTNTNDFYRGTSTVTYLN 677
QY 469 PNQ----- 471
Db 678 GSSTAQDNFTYSLGDYVWLDKKNKGVDDEKLAGVYVTLKDSNNRLEQRVTITDQSGH 737
QY 472 ---YKVEFNTPDQITTP-----YIVVNGHI----- 495
Db 738 YQFDNLQNGTYTWEFAIPDNYTPSPANNSTNDALSDGERDGTGRKVVYVAKGTINNADMT 797
QY 496 ---DPNSKGDALRLSTLYGYSNI.IWRSKMSWDNE-- 526
Db 798 VDTGFYLTPKYVGDYVWEDTNKDGIDQDNEKIGISGVKVTLLKKNKGTITGTTTDSNGY 857
QY 527 ---VAFNNG-----SGSGDGDID-----KP- 542
Db 858 EFTGLENGDYTIEFETPGYPTTKQNSGSDGKDSNGTKITVTYVKDADNKTIDSGFYKPT 917
QY 543 --- 542
Db 918 YNLGDYVWEDTNKDGIDQDSEKIGISGVKVTLLKDKNGNAIGTTTTTASGHYQFKGLENGSY 977
QY 543 -VVEQP----- 548
Db 978 TVEFETPGYPTKANSQDITVDSNGITTTGIINGADNLTIDSGFYKTPKYSVGDYVWE 1037
QY 549 ---DEPGEIEPI-----PE- 559
Db 1038 DTNKGIDQDNEKIGISGVKVTLLKDKNGNAIGTTTTTASGHYQFKGLENGSY 1097
QY 560 -----DSD 571
Db 1098 MTQTTANSNGNDEKADGEDVRVTITDHDDFSIDNGYFDDSDSDSDSDSDSDSDSDSDA 1157

APPLICANT: Poster, Timothy J.
APPLICANT: Hook, Magnus A.O.
APPLICANT: Bidhinn, Deirdre Ni
APPLICANT: Perkins, Samuel L.
TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
FILE REFERENCE: P06283US2/BAS
CURRENT APPLICATION NUMBER: US/09/200,650E
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,815
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: 60/098,427
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 3
LENGTH: 930
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-200-650E-3

Query Match 26.3%; Score 1241.5; DB 4; Length 930;
Best Local Similarity 34.9%; Pred. No. 3.7e-71;
Matches 359; Conservative 133; Mismatches 283; Indels 255; Gaps 39;

QY 1 MNMCKK-----EKHAIRKKSIGVASVLVGTLLIGFLISSKEADASENSVTQSD 48
DB 1 MNMCKKATNRKGMIPNRLNFKSRKYSVGTASILVGTLLIFG-LSGHEAKAAEHTNGELN 59
QY 49 SASNEKSNDSSSVSAAPKTDNTNVSPTKTSNTNNGETSAVQNPAAQOETTQSSNTNATT 108
DB 60 QSKNETTA--PSENKTKPKVDSRQLK-----NTQTATADQP---KVTMSDS--ATV 104
QY 109 EETPVTEATTNTQANTATTCSSNTNABELVQNTSTNTTNDNTVSVNSPQNSTN 168
DB 105 KE-----TSSNQSP-----QATANQSTKTSNTVNDKSTTYS-----NETD 144
QY 169 AENVSTTQDTSATPSNNESSAPOSTDASNKDVVNQAVNTSAPMRAFSLAAVAADAPAA 228
DB 145 KSLNTOAKDVST-----TPKTTIKPTLRMAVNTVA-----APQQ 181
QY 229 GTDITNL--TNVTVGIDSG-----TTVYHQAGYVKLVNFGFVSPNSAVKGDTPKIT 278
DB 182 GTNVNDKVHFENIDIAIDKGVNQTGKTBFWATSDVLKANYTIDDSVKEGDTEFFK 241
QY 279 VPKEINLVGTSATAKVPIM-AGQVLANGVIDSDGN-VYTFPDYNTKDDYKATLTMP 336
DB 242 YGYFRFGSVRLPSQTNLYNAQNIITAKGIYDSTNTTFTYVYDQVTVNRGSEFQV 301
QY 337 AYIDPEN--VKKTG--NVTLATIGIGSTTANKTLVDYKGYKFNLSIKGTIDQIDKTN 391
DB 302 AFARKNATTDKTAAYKMEVTL---GNDTYSEEIIIDYG-----NKAQPLISSTNYIN 351
QY 392 N---TYRQTIYVNPSSGDNVIAPLVLTGNL-----KPNTDSNALIDQONTSIKVKVDNAAD 443
DB 352 NEDLSRNMTAYVNPQKNTYTKQTFVNTLTGYKFENPA-----KNFKIYEVTDQNG 401
QY 444 LSESYFVNPENFEDYTVSNITFPNP-----QYKV-EFNTPDQITTP 486
DB 402 FVDSFTPTDKLKDVTDFDVIYSNDNKATATVLMKGQTSNNKQIITQQVAYPNSSTD- 460
QY 487 YIVVNVGHIDPNSKGLALRSLTLYGYSNIIWRSMWDEVAFNKGSGSGDGIDK----- 541
DB 461 -----NKIDYTLTD---KTKY-----SWSNSVNVGSGSTANGDKKYNLG 500
QY 542 PVVPEQDEPGE-----IEPIEDSDSPGSDSGSDNSD-----SGSDSGSDS--- 585
DB 501 DYVWEDTNKDGKQDANEKIGKGVVILKDSNGKELDRITTTDENGKYQFTGLSNGTYSVEF 560
QY 586 -----TSDSGSDASDSDASDSDASDSDAS--DSD-----SASD-----SSDN 625
DB 561 STPAGYTPPTANVGTDDAVDSDGLTTTGVIXDADNMTLDSGFYKTPKPSLGLDYVWYDSNK 620
QY 626 DSDSDS-----DSDSD-----DSDSD-----DSDSD----- 636

DB 621 DGRKRDSTEKGIGKGVKVTQLQNEKGEVIGTTTETDENGKRYRFDNLDSGKYKVFEPAGLQT 680
QY 637 --SD 694
DB 681 GTNTTDEDDKADAGEVDVITITDHD-DFTLDNGYEEETSDSDSDSDSDSDSDSDSDSDSDSDSD 739
QY 695 SD 754
DB 740 SD 799
QY 755 SDSE 814
DB 800 SD 859
QY 815 SDSE 871
DB 860 SD 917
QY 872 VPENSPKNGT 881
DB 918 ---NNSNNGT 924

RESULT 12
US-09-200-650E-5
Sequence 5 Application US/09200650E
Patent No. 6680195
GENERAL INFORMATION:
APPLICANT: Patti, Joseph M.
APPLICANT: Foster, Timothy J.
APPLICANT: Hook, Magnus A.O.
APPLICANT: Bidhinn, Deirdre Ni
APPLICANT: Perkins, Samuel L.
TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
FILE REFERENCE: P06283US2/BAS
CURRENT APPLICATION NUMBER: US/09/200,650E
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,815
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: 60/098,427
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 5
LENGTH: 1315
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-200-650E-5

Query Match 23.5%; Score 1110.5; DB 4; Length 1315;
Best Local Similarity 27.0%; Pred. No. 1.3e-62;
Matches 359; Conservative 175; Mismatches 366; Indels 431; Gaps 42;

QY 1 MNMCKKHAIRKKSIGVASVLVGTLLIGFLG--LSSKEADASENSVTQSDASNESKSD 58
DB 15 MVSNRLNKEFSIRKVTGTASILVGTLLIFGLGQKAKAESTNKELNEATTSASDNQSD 74
QY 59 SSSVSAAPKTDOT-NVSDTKTSNTNNGETSAVQNPAAQOETTQSSNTNATTEPTVTEA 117
DB 75 KVDNQQLNQEDNTKNDNQKEMVSSQGNNTTSNGKLIKESVQSTTGKV-----EV 126
QY 118 TTTTTCQANTPATQSSNTNABELVNTSNETTNDNTVSSV---NSPQNSTNAENVST 174
DB 127 STAKSDQASPKSTNEDLNTKQTSNOEALQDPLQENKSVNVQPTNEENKKVDKATEST 186
QY 175 TQDTSATEPSNNEA---PQSTDAENKDVAVQAVNTSAP-----RMRAFSLAAVAADA 225
DB 187 TLNVKSDALKSNDETLVDNNSNNENNADI--LPKSTAPKRLNTRMRIAIAVQPSSTEA 244
QY 226 PAAGTDTNOLTNVTVGIDSGTTVYHQAGYVKLVNFGFVSPNSAVKGDTPKITVPEKLN 285
DB 245 KNVNDLITSTTLTVVDADKNNKIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQV 303

US-09-134-001C-4463

Query Match 17.9%; Score 847.5; DB 3; Length 2137;

Best Local Similarity 32.0%; Pred. No. 1.4e-45;

2200 Local Similarity 32.00; Acc: 100.00; Gaps 21;
 Matches 297; Conservative 149; Mismatches 352; Indels 131; Gaps 21;

QY	33	SSKZADASNSVTQ-----SDSASNESKNDSSSVSAA-PKYDDTNVSDTK	77
Db	1076	SESEDSTSTSLSESTSTSLSGTSASTSASTSDSTSESSTSLSESTSLSDST	1135
QY	78	TSSNTNGETSVAQ--NPACQETTCQSSTNAITTEFPVTCERATTTTNCANTPATQS-	133
Db	1136	SASTSEGASTSTSESNASTSUGSLUSTISDSTSTSDSASTSTSESDSTSTSL	1195
QY	134	SNTMAEELVNOTSRETFPNDTNTVSSVNSPONSTN-AENVSTTODTTEATPPGNNESAPQ	192
Db	1196	SESTSTSLSDSTSTSTSESASTSTSESDSTSESTSLSESTSTSVSDSTA-STSDAST	1253
QY	193	STDASNKDQVVNOAVNTSAPRMARFSLAADAAPAQT---DITNQLTNTVTVIGDSITT	249
Db	1254	STSVSDGESASTSISELSTSVSDSTSTSDSASTSTSESDSTSESTSLSEST---	1310
QY	250	YPHQAGVVKLVNGFSVPNSAKVGDTFKITVPKEINLNGVTSKTAKVPIMAGDOVLANGVI	309
Db	1311	-----VSDTSASTSDSASTSTSESDSASTSLSGSTST-----LSDSTS	1352
QY	310	DSDGNVIYFTDYVNTKKDKVKATL-----TMPAYIDPENVKKTGNVTLATIGISTTANKT	364
Db	1353	TSTDASTST--SESDSERASTSLSGSTSTSLSDSTSTSDSASTSTSVSDSNSAST	1409
QY	365	VLDVEKYGKPYNLISIKGTIDQIDKNTNVKQTIYNPSPGONVIAPVLTGNLK---PNTD	421
Db	1410	SL-----SGSLSTSVSDSTSTSTSDSASASTSESDSERASTSLSGSTSTSTSDST	1459
QY	422	SNALLDQONTSIKYKVDN-----AADLSFYFNPEPFEDVTNSVNIFFPNPQY	472
Db	1460	STSTDASTSTSVSESNSTSTSEGLSTSVSDSTSTSDSASTSTSVSDSASTSS	1519
QY	473	KVFENTPDDQITTPYIVVNGHDPNKGDILALASTLYGYNSNIWRSMSWDNEAFNNG	532
Db	1520	SESVSTSDSEST-----TSTDASTST-----SVSESNSTSTSL	1556
QY	533	SGSGDGXDKVPVPEQPDEBGEIPIPDSDSDPGSDGSDGSDGSDGSDGSDGSDGSD	592
Db	1557	GSTSTSV-----SDSTSTSTSDSASTSESDSDSASTSSSESVST	1597
QY	593	SASPSDASDSDASD	652
Db	1598	SVSDTSASTSESASTSTSVSDSNSASTSLESTSTSLSDSTMSTSDSASTSTSESDSD	1657
QY	653	SD	712
Db	1658	SASTSLSDSTSTSVSESTSTSTSTSVASNSTSTSLSDSRSTSLSDSTSTSESGSTST	1717
QY	713	SD	772
Db	1718	SESDSDASTSLSESTSTSIDSTSTSTSDSASTSMVSVDSDNRASTSLSDSTSTSVSDST	1777
QY	773	SD	830
Db	1778	SASTSEGASTSTRESESTASTSLESTSTSVSDSTSTSTSDSASTSTSESDNSSETSL	1837
QY	831	SDSDSDSASTSD	888
Db	1838	SESTSTSVSDSTSTASTSASTSTSTSVSDSNSASTSLSGSTSTSVSDSTSTSCASTS	1896
QY	889	AKOSKEPLPDTGSEDEANTSIIWGLLASTI	917
Db	1897	TSES-----DSDSASTSLSGSTSTSI	1917

Search completed: November 3, 2004, 19:11:49
Job time : 56.7559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:53:31 ; Search time 193.016 Seconds
(without alignments)
2781.239 Million cell updates/sec
Title: US-09-679-643-2
Perfect score: 4725
Sequence: 1 MNMKKEKHAIRKKSIGVAS.....LASIGSLLPRKKENKDKK 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Uniprot.02:.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4725	100.0	933	2	Q53653	Q53653 staphylococ
2	4553.5	96.4	946	2	Q8XJ1	Q8xj1 staphylococ
3	4542.5	96.1	928	2	Q6GB45	Q6gb45 staphylococ
4	4430	93.8	1029	2	Q6GK4	Q6gk4 staphylococ
5	4399	93.1	989	2	Q99VJ4	Q99vj4 staphylococ
6	4359	92.3	935	2	Q32C5	Q32c5 staphylococ
7	2495	52.8	881	2	Q93MH7	Q93mh7 staphylococ
8	1810.5	38.3	913	2	Q86476	Q86476 staphylococ
9	1748.5	37.0	907	2	Q8NUL0	Q8nul0 staphylococ
10	1740.5	36.8	905	2	Q6G644	Q6g644 staphylococ
11	1638.5	34.7	877	2	Q99R07	Q99r07 staphylococ
12	1638.5	34.7	877	2	Q7A382	Q7a382 staphylococ
13	1628	34.1	1092	2	Q70022	Q70022 staphylococ
14	1610.5	34.1	1633	2	Q8CMP4	Q8cmp4 staphylococ
15	1584.5	33.5	1733	2	Q9K114	Q9k114 staphylococ
16	1559.5	32.0	873	2	Q6GDH2	Q6gdh2 staphylococ
17	1523.5	32.2	3360	2	Q8XB6	Q8xb6 lactobacill
18	1487	31.5	1056	2	Q8CQ72	Q8cq72 staphylococ
19	1423.5	30.1	1171	2	Q9KWX6	Q9kwx6 staphylococ
20	1414	29.9	1166	2	Q86489	Q86489 staphylococ
21	1395.5	29.5	1385	2	Q99W47	Q99w47 staphylococ
22	1395.5	29.5	1385	2	Q7A780	Q7a780 staphylococ
23	1356.5	28.7	957	2	Q6GBS6	Q6gbs6 staphylococ
24	1356	28.7	1637	1	PLS STAAU	P80544 staphylococ
25	1354.5	28.7	953	2	Q95W48	Q95w48 staphylococ
26	1354.5	28.7	953	2	Q7A781	Q7a781 staphylococ
27	1345	28.5	710	2	Q6R676	Q6r676 staphylococ
28	1345	28.5	710	2	AA500715	AA500715 staphylococ
29	1343.5	28.4	955	2	Q8NXX7	Q8nxx7 staphylococ
30	1325.5	28.1	1365	2	Q6GBS5	Q6gbs5 staphylococ
31	1320.5	27.9	1141	2	Q95W46	Q95w46 staphylococ

32	1320.5	27.9	1141	2	Q932F7	Q932f7 staphylococ
33	1314.5	27.8	1141	2	Q6GBS4	Q6gbs4 staphylococ
34	1314.5	27.8	1141	2	Q8NXX5	Q8nxx5 staphylococ
35	1303	27.6	1893	2	Q8KWM1	Q8kwm1 staphylococ
36	1297.5	27.5	1137	2	Q6GJA6	Q6gja6 staphylococ
37	1286.5	27.2	947	2	Q86487	Q86487 staphylococ
38	1252.5	26.5	1347	2	Q8NXX6	Q8nxx6 staphylococ
39	1155	24.4	906	2	Q6GJA7	Q6gja7 staphylococ
40	1110.5	23.5	1315	2	Q86488	Q86488 staphylococ
41	1055.5	22.3	1698	2	Q8LC00	Q8lc00 staphylococ
42	963	20.4	903	2	Q8Y697	Q8y697 listeria mo
43	957.5	20.3	882	2	Q92AK9	Q92ak9 listeria in
44	883.5	18.7	1253	1	DSPP HUMAN	Q9nzw4 homo sapien
45	880	18.6	931	2	Q9K113	Q9k113 staphylococ

ALIGNMENTS

RESULT 1
Q53653 PRELIMINARY; PRT; 933 AA.
ID Q53653
AC Q53653;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RA MEDLINE=9422142; PubMed=8170386;
RA McDevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrogen receptor
of Staphylococcus aureus.";
RL Mol. Microbiol. 11:237-248(1994)
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (by similarity).
DR EMBL; Z18852; CAA79304.1; -.
DR PIR; S41539; S41539.
DR PDB; 1N67; X-ray; A=202-560.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 933 AA; 97057 MW; EB51A6DE2FF759F4 CRC64;

Query Match	100.0%;	Score 4725;	DB 2;	Length 933;
Best Local Similarity	100.0%;	Pred. No. 3.3e-194;		
Matches 933;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MNMKKKEKHAIRKKSIGVASVLVGLTIGFGLLSKEADASENSVTQSDASNESKNDSS	60	
Db	1	MNMKKKEKHAIRKKSIGVASVLVGLTIGFGLLSKEADASENSVTQSDASNESKNDSS	60	
Qy	61	SVSAPKTDITNVSDTKTSSNTNNGETSVAPQAQOETQSSSTNATTEEPVTGEATTT	120	
Db	61	SVSAPKTDITNVSDTKTSSNTNNGETSVAPQAQOETQSSSTNATTEEPVTGEATTT	120	
Qy	121	TTNQANTPATQSSNTNAEELVNQTSNETTENDNTVSSVNSPQNSTNAEVNSTQDTST	180	
Db	121	TTNQANTPATQSSNTNAEELVNQTSNETTENDNTVSSVNSPQNSTNAEVNSTQDTST	180	
Qy	181	EATPSNNEAPQSDASKDKVNVQAVNTSAPRMAFSLAAVAADAPAGTDTTNQLTNVT	240	

Db 181 EATPSNNEAPQSTDASNKDVVNQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLTNVT 240
QY 241 VGI DSGTIVYPHOAGVYKLYGFSVPNSAVKGDTEFKITVPKELNLNGVSTAKVPPIMAG 300
Db 241 VGI DSGTIVYPHOAGVYKLYGFSVPNSAVKGDTEFKITVPKELNLNGVSTAKVPPIMAG 300
QY 301 DQVLANGVIDSDGNVIYFTDYVNTKDDVKATILMPAYIDPENVKKTGNVTLATGIGSTT 360
Db 301 DQVLANGVIDSDGNVIYFTDYVNTKDDVKATILMPAYIDPENVKKTGNVTLATGIGSTT 360
QY 361 ANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYROTIIYVNPNGDNVIAVLTKNLKPNT 420
Db 361 ANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYROTIIYVNPNGDNVIAVLTKNLKPNT 420
QY 421 DSNALIDQQNTSIKYVKVDNAADLSSEYFVNPENFEDVTNSVNI TFPNPNQYKVEFTPD 480
Db 421 DSNALIDQQNTSIKYVKVDNAADLSSEYFVNPENFEDVTNSVNI TFPNPNQYKVEFTPD 480
QY 481 DQITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNI IWRMSMWDNEVAFNNGSGSDGID 540
Db 481 DQITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNI IWRMSMWDNEVAFNNGSGSDGID 540
QY 541 KPVPPEQDPPEGEIEPIPEDSDPGSDGSDNSDSDSGSDSTSDSGSDSASDSDSA 600
Db 541 KPVPPEQDPPEGEIEPIPEDSDPGSDGSDNSDSDSGSDSTSDSGSDSASDSDSA 600
QY 601 SDSDSASDSDSASD 660
Db 601 SDSDSASDSDSASD 660
QY 661 SD 720
Db 661 SD 720
QY 721 SD 780
Db 721 SD 780
QY 781 SD 840
Db 781 SD 840
QY 841 SDGSDSDSDSDSDSDSDSDSDSDSGSNVVPNPNPKNGTNAKNKNEAKDSKEPLPDTG 900
Db 841 SDGSDSDSDSDSDSDSDSDSDSDSGSNVVPNPNPKNGTNAKNKNEAKDSKEPLPDTG 900
QY 901 SEDEANTSLIWGLASIGSLLLFRKKENKDKK 933
Db 901 SEDEANTSLIWGLASIGSLLLFRKKENKDKK 933

RESULT 2

Q8NXJ1 PRELIMINARY; PRT; 946 AA.
AC Q8NXJ1; 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Name=clfa; OrderedlocusNames=MW0764;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwana N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827 (2002).

CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
CC EMBL; AF004824; BAB94629.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 946 AA; 98237 MW; EFPB838793201173 CRC64;

Query Match 96.4%; Score 4553.5; DB 2; Length 946;
Best Local Similarity 95.4%; Pred. No. 7.2e-187;
Matches 903; Conservative 15; Mismatches 14; Indels 15; Gaps 2;

QY 1 MMKKKEKHAIRKKSIGVASVLVGLTIGFLLSSKEADASENSVTQDSASNEKSNDS 60
Db 1 MMKKKEKHAIRKKSIGVASVLVGLTIGFLLSSKEADASENSVTQDSASNEKSNDS 60
QY 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVQNPAAQOETTQSSSTNATTEETPTVGEATT 120
Db 61 SVSAAPKTDITNVSDTKTSSNTNNGETSVQNPAAQOETTQSSSTNATTEETPTVGEA-TT 119
QY 121 TTQANTPATTOSSNTNABELVNTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDTST 180
Db 120 ATQANTPATTOSSNTNABELVNTSNETTSDNTNTVSSVNSPQNSTNAENVSTTQDTST 179
QY 181 EATPSNNEAPQSTDASNKDVVNQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLTNVT 240
Db 180 EATPSNNEAPQSTDASNKDVVNQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLTNVT 239
QY 241 VGI DSGTIVYPHOAGVYKLYGFSVPNSAVKGDTEFKITVPKELNLNGVSTAKVPPIMAG 300
Db 240 VGI DSGTIVYPHOAGVYKLYGFSVPNSAVKGDTEFKITVPKELNLNGVSTAKVPPIMAG 299
QY 301 DQVLANGVIDSDGNVIYFTDYVNTKDDVKATILMPAYIDPENVKKTGNVTLATGIGSTT 360
Db 300 DQVLANGVIDSDGNVIYFTDYVNTKENVTANITMPAYIDPENVTKTGNVTLTTGIGSTT 359
QY 361 ANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAVLTKNLKPNT 420
Db 360 ANKTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAVLTKNLKPNT 419
QY 421 DSNALIDQQNTSIKYVKVDNAADLSSEYFVNPENFEDVTNSVNI TFPNPNQYKVEFTPD 480
Db 420 DSNALIDQQNTSIKYVKVDNAADLSSEYFVNPENFEDVTNSVNI TFPNPNQYKVEFTPD 479
QY 481 DQITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNI IWRMSMWDNEVAFNNGSGSDGID 540
Db 480 DQITTPYIVVNGHIDPNKSGDLALRSTLYGYDSRPFWRMSMWDNEVAFNNGSGSDGID 539
QY 541 KPVPPEQDPPEGEIEPIPEDSDPGSDGSDNSDSDSGSDSTSDSGSDSASDSDSA 600
Db 540 KPVPPEQDPPEGEIEPIPEDSDPGSDGSDNSDSDSGSDSTSDSGSDSASDSDSA 599
QY 601 SDSDSASDSDSASD 646
Db 600 SDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDSDSDSDSDSDSDSDSD 659
QY 647 SD 706
Db 660 SD 719
QY 707 SD 766
Db 720 SD 779
QY 767 SASD 826

RT evidence for the rapid evolution of virulence and drug resistance. ;
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).

DR EMBL; BX571856; CAG39851.1; -.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; LPTXG_anchor; 1.
DR TIGRFAMs; TIGR01167; Y5IRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 1029 AA; 106715 MW; 025E13C6A4C4F020 CRC64;

Query Match 93.8%; Score 4430; DB 2; Length 1029;
Best Local Similarity 86.3%; Pred. No. 1.5e-181;
Matches 888; Conservative 27; Mismatches 18; Indels 96; Gaps 5;

QY 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSNDS 60
DB 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSNDS 60
QY 61 SVSAAPKTDNTNVSDDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120
DB 61 SVNAAPKTDNTNVSDDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120
QY 121 TTQANTPATTOSSNTAELVNOTSNETTNDTNTVSSVNSPQNTAENVTSTQDST 180
DB 121 ATQANTPATTOSSNTAELVNOTSNETTNDTNTVSSVNSPQNTAENVTSTQDST 180
QY 181 EATPSNNEGAPOSTDASNDKVVNOAVNTSAPRMRAFSLAAVAADAPAAAGDTITNLTNTV 240
DB 181 EATPSNNEGAPOSTDASNDKVVNOAVNTSAPRMRAFSLAAVAADAPAAAGDTITNLTNTV 240
QY 241 VGIDSGTTPYHQAGYVKLVNFGFSPNSAVKGDFTFKITVPKELNLNGVTSSTAKVPIMAG 300
DB 241 VGIDSGTTPYHQAGYVKLVNFGFSPNSAVKGDFTFKITVPKELNLNGVTSSTAKVPIMAG 300
QY 301 DOVLANGVIDSGNVIYTFDVTYKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTT 360
DB 301 DOVLANGVIDSGNVIYTFDVTYKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTT 360
QY 361 ANKTVLVDEYKYGKFNLSIKGTIDDKTNNYROTIIYVNSGDNVIAPIVLTGNLKPNT 420
DB 361 ANKTVLVDEYKYGKFNLSIKGTIDDKTNNYROTIIYVNSGDNVIAPIVLTGNLKPNT 420
QY 421 DSNALIDQONTSIKVKVVDNAADLGSFYFNPENFEDVTNSVNIITPPNPQYKVFENTPD 480
DB 421 DSNALIDQONTSIKVKVVDNAADLGSFYFNPENFEDVTNSVNIITPPNPQYKVFENTPD 480
QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYNLSNIIFWSMWDNEVAFNNGSGGDGID 540
DB 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYNLSNIIFWSMWDNEVAFNNGSGGDGID 540
QY 541 KPVPPEQDPEGEIPIPEDSDSDP-----GSDSGSDNSDSDSGSDSTSDSGSDSAD 596
DB 541 KPVPPEQDPEGEIPIPEDSDSDP-----GSDSGSDNSDSDSGSDSTSDSGSDSAD 600
QY 597 SDSASDSASDSASDSDA-----SDSDSASDSASDSND 626
DB 601 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 660
QY 627 SPDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 686
DB 661 SPDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 720
QY 687 SPDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 732
DB 721 SPDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 780
QY 733 SPDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 791

DB 781 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 840
QY 792 -----SDSDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 846
DB 841 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 900
QY 847 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 864
DB 901 SDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 960
QY 865 SGNNNNVVPNPKNGTNAKNEAKDSKEPLPDTGSEDEANTSLIWLASIGSLILFR 924
DB 961 SGNNNNVVPNPKNGTNAKNEAKDSKEPLPDTGSEDEANTSLIWLASIGSLILFR 1020
QY 925 RKKNKDKK 933
DB 1021 RKKNKDKK 1029

RESULT 5

Q99VJ4 PRELIMINARY; PRT; 989 AA.
ID Q99VJ4
AC Q99VJ4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein A, clumping factor.
GN Name=clfa; OrderedLocNames=SA0742;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]_TaxID=158879;
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaiko C.,
RA Sekimizu K., Hirakawa H., Kuhnara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus".
RL Lancet 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
DR EMBL; AP003131; BAB41975.1; -.
DR PIR; D89852; D89852.
DR HSP; 053653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPTXG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

Query Match 93.1%; Score 4399; DB 2; Length 989;
Best Local Similarity 88.6%; Pred. No. 3.1e-180;
Matches 876; Conservative 27; Mismatches 30; Indels 56; Gaps 2;
QY 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSNDS 60
DB 1 MNMKKEKHAIKKKSGVSLVGLTIGFLLSSKEADASENSVTQSDASNEKSNDS 60
QY 61 SVSAAPKTDNTNVSDDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120
DB 61 SVSAAPKTDNTNVSDDTKTSNTNNGETSVQAQPAQOETTQSSSTNATTEETPTVGEATT 120

QY 121 TTNOANTPATTOSSNTNAELVNQTSNETTNDNTNTVSSVNSPONSNAENSVTQDTST 180
Db 121 TTNOANTPATTOSSNTNAELVNQTSNETTNDNTNTVSSVNSPONSNAENSVTQDTST 180
QY 181 EATPSNNEAPQSTASNDKDVQAVNTSAPRMRFAFLAAVAADAPAAAGTDTITNQLTNVT 240
Db 181 EATPSNNEAPQSTASNDKDVQAVNTSAPRMRFAFLAAVAADAPAAAGTDTITNQLTNVT 240
QY 241 VGDSTGTVPHQAGYVVKLVGFSVPNSAVKGDFTFKITVPKELNLNGVTSTAKVPPIMAG 300
Db 241 VGDSTGTVPHQAGYVVKLVGFSVPNSAVKGDFTFKITVPKELNLNGVTSTAKVPPIMAG 300
QY 301 DOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTPAYIDPENVKTKGNVTLATIGSTT 360
Db 301 DOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTPAYIDPENVKTKGNVTLATIGSTT 360
QY 361 ANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAVPLTGNLKPNT 420
Db 361 ASKTVLIDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAVPLTGNLKPNT 420
QY 421 DSNALIDQOQNTSIKVKYKVDNAALSESFYVNPENFEDVTNSVNTFPNPQYKVEFTPD 480
Db 421 KSNALIDAKNTDIKVRVDNANDLSSEYVNPSPDFEDVTNQVIRSFNNANQYKVEFTPD 480
QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
Db 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
QY 541 KPVPVEQDFPGEIETIPEDSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
Db 541 KPVPVEQDFPGEIETIPEDSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
QY 601 SDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDS 622
Db 601 SDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDSASDSDS 660
QY 623 -----SDNPSD 676
Db 661 SDSDSASD 720
QY 677 SD 736
Db 721 SD 780
QY 737 SD 796
Db 781 SD 840
QY 797 SD 853
Db 841 SD 900
QY 854 -----DSESD 904
Db 901 STDTGSDNDSDSNDSSESGSNVPPNPNKNGTNASKNKAKDSKEPLPTGSEDE 960
QY 905 ANTSLIWGLASIGSLILFRKKENKOKK 933
Db 961 ANTSLIWGLASIGSLILFRKKENKOKK 989
RESULT 6
Q932C5
ID Q932C5 PRELIMINARY; PRT; 935 AA.
AC Q932C5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Name=fnb; OrderedLocusNames=SAV0811;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

NCBI_TaxID=158878;
[1]
SEQUENCE FROM N.A.
STRAIN=Mu50 / ATCC 700699;
MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
Mizufani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Shiba T.,
Kanehisa M., Yamashita A., Oshino K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus";
RL Lancet 357:1225-1240 (2001).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
CC EMBL; AP003360; BAB56973.1; -.
DR HSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRfams; TIGR01167; LpXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 935 AA; 96949 MW; DC5A2D92CE3BA91C CRC64;
Query Match 92.3%; Score 4359; DB 2; Length 935;
Best Local Similarity 92.5%; Pred. No. 1.5e-178;
Matches 865; Conservative 28; Mismatches 40; Indels 2; Gaps 1;
QY 1 MNMKKEKHAIKKSIGVASVLVGLTIGFGLLSKEADASENSVTQSDSASNESKSDSS 60
Db 1 MNMKKEKHAIKKSIGVASVLVGLTIGFGLLSKEADASENSVTQSDSASNESKSDSS 60
QY 61 SVSAAKPTDNTNVDTKTSNTNNGETSVACNPAQQTQSSSTNATTEPTVTGEATTT 120
Db 61 SVSAAKPTDNTNVDTKTSNTNNGETSVACNPAQQTQSSSTNATTEPTVTGEATTT 120
QY 121 TTNOANTPATTOSSNTNAELVNQTSNETTNDNTNTVSSVNSPONSNAENSVTQDTST 180
Db 121 TTNOANTPATTOSSNTNAELVNQTSNETTNDNTNTVSSVNSPONSNAENSVTQDTST 180
QY 181 EATPSNNEAPQSTASNDKDVQAVNTSAPRMRFAFLAAVAADAPAAAGTDTITNQLTNVT 240
Db 181 EATPSNNEAPQSTASNDKDVQAVNTSAPRMRFAFLAAVAADAPAAAGTDTITNQLTNVT 240
QY 241 VGDSTGTVPHQAGYVVKLVGFSVPNSAVKGDFTFKITVPKELNLNGVTSTAKVPPIMAG 300
Db 241 VGDSTGTVPHQAGYVVKLVGFSVPNSAVKGDFTFKITVPKELNLNGVTSTAKVPPIMAG 300
QY 301 DOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTPAYIDPENVKTKGNVTLATIGSTT 360
Db 301 DOVLANGVIDSDGNVIYFTDYVNTKDDVKATLTPAYIDPENVKTKGNVTLATIGSTT 360
QY 361 ANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAVPLTGNLKPNT 420
Db 361 ASKTVLIDYKYGKFNLSIKGTIDQIDKTNNTYQTIYVNPNGDNVIAVPLTGNLKPNT 420
QY 421 DSNALIDQOQNTSIKVKYKVDNAALSESFYVNPENFEDVTNSVNTFPNPQYKVEFTPD 480
Db 421 KSNALIDAKNTDIKVRVDNANDLSSEYVNPSPDFEDVTNQVIRSFNNANQYKVEFTPD 480
QY 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
Db 481 DOITTPYIVVNGHIDPNSKGLALRSTLYGYSNIIWRSMWMDNEVAFNNGSGSGDID 540
QY 541 KPVPVEQDFPGEIETIPEDSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600
Db 541 KPVPVEQDFPGEIETIPEDSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSD 600

[illegible]

```
Db 69 IETPOLNTTANDTSDISANTNSANVDSTAKPMSTQTSNTTTEPASTNETP-----QLTAI 124
QY 122 TNOANTPATTSQSNNTNABELVQTSNETTNDNTNTVSSVNSPQNSNAENVSTQDTST 181
Db 125 KQOAT--AAKMQDQVTPQANSQVDNKT--NDANSI-----ATNSE-----LKNPQL 169
QY 182 ATPSNNEAPQSTDSASNDKDVNOAVNTSAPRMEAFSLAAVAADAP-----AAGDTITNQ 235
Db 170 DLP---QSSPQT-----ISNAQGTSPKSVYTRAVRSLAAVAPVNAADAKGTINVDK 218
QY 236 LTNVTVGIDSGTTPVPHOAGYVKNLNGFSPNSAVKGDTEKTVPKELNLNGVTSTAKVP 295
Db 219 VTAKDFQLEK--TTFDPNQSNTFMAANFTVTGQVKSQDYF-----TAKUP 262
QY 296 PIMAGDQVLANGVID-----SDGNVI-----YTFTDYVNTKD 327
Db 263 -----DSVTGNGDVDSYNSNNMTPIADIIVDKNEVAKATVDILTKTYTFTVDYVNDKQ 317
QY 328 DVKATLTWPAYIDPENVKTKTNVLTATIGISTTANKTVLVLDYEK-----YGFYNLSIKGT 383
Db 318 NINGKFSPLFTDRAKAPKSGTYDANINIADMFNNKITYNSSPIAGIDKPGANISSQ 377
QY 384 IDQIDKTN--NTYRQTIYVNP--ENFEDVTSV--NITFPNPQYKVFNPDPDQITTPYVWVNG 493
Db 378 IIGVDTASQNTYKQTVFVNPQRVLGNTWV--YIKGYQDKIESSGKVSATDTKLRIPE 435
QY 438 VDMAADLSSEYFVNP--ENFEDVTSV--NITFPNPQYKVFNPDPDQITTPYVWVNG 493
Db 436 VNDTSKLSYVADPNSNLKEVTQFKDKITYKQNVASINFG-----DINKTYVVLVEG 491
QY 494 HIDPNSKGDALRSTLYGYN-----SNIIWRSMWNEVAFNNGSGDGIKPVVPEQPD 549
Db 492 HYDKTGKN---LKTQVIQENVDPATGKDYISFGWNNENVRYGGGSADG--DSAVNPKDP 547
QY 550 EFGETIEPIE--DSDSDPGSDSGSDNSDSDSGSDSTSDSGSDSASDSDSASDSD 608
Db 548 PGPPVPPEPPEPPEPDPPEPTPPPEPDPPEPDPPEPDPPEPDPPEPDPPEPDP 607
QY 609 SDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 668
Db 608 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 667
QY 669 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 728
Db 668 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 727
QY 729 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 788
Db 728 SESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 787
QY 789 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 848
Db 788 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 834
QY 849 SSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 907
Db 835 -----RVTPPNEQAPSPNPKGEVNSKVSQKHQTDALPETGDKSENTNA 880
QY 908 SLIWLGLLASIGSLLLFREKKENKOKK 933
Db 881 TLFGMWALLGSLLLFRKKQDQHK 906
RESULT 10
Q6G644
TID Q6G644 PRELIMINARY; PRT; 905 AA.
AC Q6G644;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Fibrinogen and keratin-10 binding surface anchored protein.
GN ORNameS=SAS2516;
OS Staphylococcus aureus subsp. aureus MSSA476.
```

```
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=MSSA476;
RC RA Holden M.T.G., Feil R.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Fellwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Leonard N., Line A., Mayes R., Moulé S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; BX571857; CAG44333.1; -.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; LPXTG_signal; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 905 AA; 96411 MW; B8B341A39D800DB8 CRC64;

Query Match 36.8%; Score 1740.5; DB 2; Length 905;
Best Local Similarity 42.5%; Pred. No. 1.3e-66;
Matches 419; Conservative 142; Mismatches 276; Indels 149; Gaps 31;

QY 5 KKEKHAIRKKSIGVASVLVGLIGLLSSKEADASE--NSVTQSDSASNEKSDSSSV 62
Db 11 KONKYSIRRFVTGTTTSVIGATILFG-ICNHOQAQASEQSDNTTQS-SKNNASADSEKNNM 68
QY 63 SAAPKTDITNTSDTKTSNTNGETSVAGNPAQOETQSSSTN-ATTETPTVGTATTT 121
Db 69 IETPOLNTTANDTSDISANTNSANVDSTAKPMSTQTSNTTTEPASTNETP-----QLTAI 124
QY 122 TNOANTPATTSQSNNTNABELVQTSNETTNDNTNTVSSVNSPQNSNAENVSTQDTST 181
Db 125 KQOAT--AAKMQDQVTPQANSQVDNKT--NDANSI-----ATNSE-----LKNPQL 169
QY 182 ATPSNNEAPQSTDSASNDKDVNOAVNTSAPRMEAFSLAAVAADAP-----AAGDTITNQ 235
Db 170 DLP---QSSPQT-----ISNAQGTSPKSVYTRAVRSLAAVAPVNAADAKGTINVDK 218
QY 236 LTNVTVGIDSGTTPVPHOAGYVKNLNGFSPNSAVKGDTEKTVPKELNLNGVTSTAKVP 295
Db 219 VTAKDFQLEK--TTFDPNQSNTFMAANFTVTGQVKSQDYF-----TAKUP 262
QY 296 PIMAGDQVLANGVID-----SDGNVI-----YTFTDYVNTKD 327
Db 263 -----DSVTGNGDVDSYNSNNMTPIADIIVDKNEVAKATVDILTKTYTFTVDYVNDKQ 317
QY 328 DVKATLTWPAYIDPENVKTKTNVLTATIGISTTANKTVLVLDYEK-----YGFYNLSIKGT 383
Db 318 NINGKFSPLFTDRAKAPKSGTYDANINIADMFNNKITYNSSPIAGIDKPGANISSQ 377
QY 384 IDQIDKTN--NTYRQTIYVNP--ENFEDVTSV--NITFPNPQYKVFNPDPDQITTPYVWVNG 493
Db 378 IIGVDTASQNTYKQTVFVNPQRVLGNTWV--YIKGYQDKIESSGKVSATDTKLRIPE 435
QY 438 VDMAADLSSEYFVNP--ENFEDVTSV--NITFPNPQYKVFNPDPDQITTPYVWVNG 493
Db 436 VNDTSKLSYVADPNSNLKEVTQFKDKITYKQNVASINFG-----DINKTYVVLVEG 491
QY 494 HIDPNSKGDALRSTLYGYN-----SNIIWRSMWNEVAFNNGSGDGIKPVVPEQPD 549
Db 492 HYDKTGKN---LKTQVIQENVDPATGKDYISFGWNNENVRYGGGSADG--DSAVNPKDP 547
```


[illegible][illegible]

Query Match	34.1%	Score 1610.5	DB 2	Length 1633
Best Local Similarity	29.5%	Pred. No. 9.4e-61		
Matches 480	Conservative 153	Mismatches 282	Indels 713	Gaps 39
Qy	5	KKEKHARKKSGIVASVLVGLTGFGLISKEADASENSVTQSDASNE	---	SKSNDSS 60
Db	19	KVKYSIRKFTVGTASILVGLATLMFG-AADNEAKAAEDN--QLESASKEQKGRSDNESS	75	
Qy	61	SVS-----AAPKTD-----TNVS-DTKTSNT	82	
Db	76	KLNQVDLDNGSHSSEKTTNVNNAEVEKKVPEPTTSDVSKPKANEAVTNESKPKTTEAP	135	
Qy	83	NGTGSVAQNP-----AQETTQSS-----STNAATEETPTVGTBATT	---	TTTQCAN 126
Db	136	TWNEESIAETPKTSTTQODSTPEKNKPSLKDNLNSSTTSKESKTDDEHSTKQAOQWSTKN	195	
Qy	127	--TPAATQSSNTNAEEVLNQTSE-----TTFND-----	---	153

```
Db 196 LDTNDSPTQSEKTSQANDSTDNQAPSKQLDSKPSQKVYKTKFNDSEPTQDVEHTTK 255
QY 154 -----TNTVSSVNSQNSNAENVSTQDTS-----TEATPSNNESAPOSTDASNKDVVN 203
Db 256 LKTPSISTDSSVNDKQDYTRSAVSLGVDSNETEAITNAVRDNLDLKAASREQINEAIIA 315
QY 204 Q-----AVNTS-----APRMRAFSLAAVAADAAPAAAGTDITN-- 234
Db 316 EALKXDFSNPDYGVDTPLALNTSQSKNSPHKSPASPRMLMSLAA-----EPNSGKVNDKV 371
QY 235 QLTNVTVGID-----SGTTVYPHQAGYVKLANTGVSPNSAVKGDFTFKITVPKELNLNGVT 289
Db 372 KITNPTLSLNKSNHANNVIMPTSENEQFNLKANYELDDSIKEGDTFTIKYQYIRPGGLE 431
QY 290 STAKVPPIMAGO-QVLANGVIDSGN-VLYTETDYNTKDYKAILTMEAYIDPEN-VKK 346
Db 432 LPAITKQURSKDGSIVANGVYDKTNTTFTYTFNVDYQONITGSPDLIATPKRETAIKD 491
QY 347 TGNVTLATIGSTTANKTVLDVYKGYKFNLSIKGTIDQIDKTNNTYRQTIYVNPSPGN 406
Db 492 NQNYPMEVITIANEVVKOFIVD---YGNKKDNTTTAAVANVDNVNKKHNEVVYLNQNNQ 548
QY 407 -----VIAPVLGNLKP-----TDSNALIDQONTSI----- 433
Db 549 PKYAKYFSTVKNKGFIPGEVKYVEYVTDNTNAMVDSFNPDLNSSNVKDVTSQFTPKVSADGT 608
QY 434 -----KVKVDNAAD-----LSSEYFVYNPFENFEDVTVNSVNIPTN 468
Db 609 RVDINPARSMANGKYIVTQAVRPTGTGNTGYTEYWLTRDGTNTNDFFYGTGKSTTVTYLN 668
QY 469 PNQ-----YKVEFTPDQITTP-----YIVVVNGHI----- 471
Db 669 GSSTAQGDNPVYSLGYYVWLDKKNKGQVDDDEKGLAGVYVVTCLKDSNNRELQRTVTDQSH 728
QY 472 -----YKVEFTPDQITTP-----YIVVVNGHI----- 495
Db 729 YQFDNLQNGTYRVEFAIPDNVTPSPANNSTDAIDSDGERDGRKRVVAKGTINNADNMT 788
QY 496 -----DPNSKGDALASTLYGYSNIIWRSMWDNE-- 526
Db 789 VUTGYLTPKYNVGVYVWEDTNKGDIQDNEKIGSNVTKLNKNGDTIGTITTDTSNGKY 848
QY 527 --VAFNNG-----SGSGDGID-----KPV 543
Db 849 EFTGLENGDYTFEFETPEGYTTPKQNSGSGDEKDSNGTKTIVTVKDAADKNTDGSFYKPI 908
QY 544 VP----- 545
Db 909 YNLGDIYVWEDTNKGDIQDDEKIGSGVKVTLKDKNGNAIGTTTTDASGHYQFKGLENGSY 968
QY 546 -----EOP----- 548
Db 969 TVEFTPSGYTPKANSQGDITVDSNGITTTGIIINGADNLITDGSFYKTPKYSVGDIYWE 1028
QY 549 -----DEPGHIEPT-----PE- 559
Db 1029 DTNKDGIQDNEKIGSGVKVTLKDEKGNLISITTTTDDENGKYQFDNLDSGNYIIHHEKPEBG 1088
QY 560 ----- 559
Db 1089 MTQTTANSNGNDEKADGEDVRVTITTDHDFSIDNGYFDDSDSDSDADSDSDSDSDA 1148
QY 560 ----- 559
Db 1149 DSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDADSDADSDSDSDSDSDSDSDS 1208
QY 560 -----DSDSDPGS 567
Db 1209 DSDSDSDSDSDSDSDSDSDSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDS 1268
QY 568 DSGSDNSDSDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 627
Db 1269 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1328
```

```
QY 628 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 687
Db 1329 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1388
QY 688 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 747
Db 1389 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDA 1448
QY 748 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 807
Db 1449 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1508
QY 808 ESDSDSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 867
Db 1509 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1568
QY 868 NNVVFPNPSKNGTNNASKNKNAKDSKEPLDPTGSEDEANT--SLIWGLIASIGSLL-PR 924
Db 1569 DSD---SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1625
QY 925 RKKNKDK 932
Db 1626 RKKNKEX 1633
```

RESULT 15

```
Q9K114 PRELIMINARY; PRT; 1733 AA.
ID Q9K114;
AC Q9K114;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative cell-surface adhesin SdrF.
GN Name=sdrF;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9491;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Bihdin D., Lina G.,
RA Speciale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
DR EMBL; AF245041; AAF72509.1; -.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; GpoS_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 4.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
SQ SEQUENCE 1733 AA; 184720 MW; D8D62EA1692FD4E8 CRC64;
```

```
Query Match 33.5%; Score 1584.5; DB 2; Length 1733;
Best Local Similarity 30.6%; Pred. No. 1.3e-59;
Matches 452; Conservative 156; Mismatches 285; Indels 583; Gaps 35;
QY 5 KKEHAIKKSIGVASVLVGLTIGFGLLSKKEADASENSVTQDSASNE---SKSDSDS 60
Db 19 KVKYSIRKFTVGTAGILVGLTLMFG-AADNEAKAEDN--QLESASKEQKSGRDNENS 75
QY 61 SVS-----AAPKTTD-----TNVS-DTKTSNT 82
Db 76 KLVQDLNGSHSSEKTTNNVNAVEVKVEAPTSDSVKPKANEAVVTNESTKPKTEAP 135
QY 83 NNGETSVAQNP-----AQOETQSS-----STNATTEETPVTEGAVT-----TTTNQAN 126
```

Db 136 TVNEESIAETPKTSTTQOOSTEKNNPSLKDNLNSSSTTSKESKTDBHSTKQAOQWSTNKS 195
Qy 127 ---TPATTQSSNTNAELVNOTSNE-----TTND----- 153
Db 196 LDTNDSPTQSEKTSQANNDSTNQAPSQKDSKSEQKVYTKFNDEPTQVEHTTK 255
Qy 154 ---TNTVSSVNSPQNSTNAENVSTTQDTS---TEATPSNNEGAPOQSTDAENKDVVN 203
Db 256 LKTPSVSTSSVNDKQDTRSAVASLGVDSNEEAITNAVRDNLDLKAAAREQINEALIA 315
Qy 204 QAV-----NTSAPRMRAFSLAAVAADAPAAAGTDITN-- 234
Db 316 EALKKDFSPNDYGVDTPLALNRQSQKNSPKHSASPRNMLSLAA---EPNSGKNVNDKV 371
Qy 235 OLNTNVLGID-----SGTIVYHQAGVVKVLYNFGFVSPNSAVKGTFFKITVPEKELNNGVT 289
Db 372 KITNPTLSLNKSNHANNVWTPTSNEOFNLKANYELDDSIKEGDTFTIKYQVIRPGGLE 431
Qy 290 STAKVPPIMAGD-QVLANGVIDSGN-VIYTFDYVNTKDDVKATLTPAYIDPEN-VKK 346
Db 432 LPAIKTQLRSKDSIVANGVYDKTNTTTFYFNVQYQNIITGSPDLIATPKRETAIKD 491
Qy 347 TGNVTLATGISTANKTVLVDYEKYGKFNLSIKGTIDQIDKTNNTYROTIVYNPBGDN 406
Db 492 NQNPMEVTTIANEVVKDFIVD---YGNKKDNTTTAAVANVDVNVNKHNEVVYLNQNNQN 548
Qy 407 -----VIAPVLTGNLKEN-----TDSNALIDQOQNTS1----- 433
Db 549 PKYAKFSTVKNGEFIPGEVKVVEVDTNAWVDSPNLDLSSNVKDVTSQFAPKVSADGT 608
Qy 434 -----KVYKVDNAAD-----LSESYFVNPNFEDVTSVNIITFPN 468
Db 609 RVDINPARSMANGKYYIVTQAVRPTGNGVYTEWLTFRDGTNTNDFYRGTKSTTVTYLN 668
Qy 469 PNQ----- 471
Db 669 GSSTAQDNPTYSLGDIYVWLDKKNKGVDDEKGLAGVYVTLKDSNNRELQRTVTTQSGH 728
Qy 472 -----YKVEFNTPDQIITP-----YIVVNVNGHI----- 495
Db 729 YQFNLQNGTYTVFAPDPNTPSPANNSTNDAIDSGERDGRKRVVAKGTINNADNMT 789
Qy 496 -----DPNSKGLALRSTLYGYNSNIIWRSMWDNE-- 526
Db 789 VDTGFFYLTPKNVGDIYWEDTNKDGIDQDNEKGLSGVKVTLKKNKGDTIGTTITDTSNGY 848
Qy 527 --VAPNNG-----SGSGDGID-----KP- 542
Db 849 EFTGLENGDYTIEFETPEGTPTKQNSGSDRGKDSNGTKTIVTVKADNKTIDSGFYKPT 908
Qy 543 ----- 542
Db 909 YNLGDIYWEDTNKDGIDQDSEKGISGVKVTLLKDKNGNAIGTTTTTDSGHYQFKLENGSY 968
Qy 543 -VVPEQF----- 548
Db 969 TVREFETSGYTPTKANSQDITVDSNGITTTGIINGADNLITDSGFYKPKYSVGDIYWE 1028
Qy 549 -----DEPGEIEPI-----PE- 559
Db 1029 DTNKGDIQDNEKGISGVKVTLLKDEKGNIIISTTTTDENGKYQFNLDPGNGYIIHFKEPEG 1088
Qy 560 -----DSDSDPFSGSGS 571
Db 1089 MTQTTANGNDDEKADGEDVRVITIDHDFSIDNGYFDDSDSDSDSDSDSDSDA 1148
Qy 572 DSNDSGSGSDSTSDSGSDSASDSASDSASDSASDSASDSASDSASDSASDSASDS 631
Db 1149 DSDSDADSD 1208
Qy 632 DSD 691
Db 1209 DSD 1268

Qy 692 DSD 751
Db 1269 DSDSDADSD 1328
Qy 752 DSD 811
Db 1329 DSD 1388
Qy 812 DSESD 871
Db 1389 DSD 1448
Qy 872 VPNSPKNGTNGNKNKNEAKDSKEPLPTGTSEDEANT 907
Db 1449 DSD 1484

Search completed: November 3, 2004, 19:10:15
Job time : 204.016 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:48:10 ; Search time 100.005 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VGLTIGFLLSSREADSEN.....NGSGSGDGDKPVVPEQDE 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2702	100.0	933	3	AA558435	Aa558435 Staphyloc
2	2702	100.0	933	4	AB69508	Ab69508 Staphyloc
3	2694	99.7	933	6	ABJ18947	Abj18947 Pathogen
4	2690	99.6	927	6	ABM72221	Abm72221 Staphyloc
5	2690	99.6	936	2	AAW89801	Aaw89801 Staphyloc
6	2621	97.0	520	5	AAE22262	Aae22262 Staphyloc
7	2466	91.3	935	6	ABU16402	Abu16402 Protein e
8	2450	90.7	1021	4	AAU33975	Aau33975 Staphyloc
9	2450	90.7	1021	4	AAU38951	Aau38951 Staphyloc
10	2408	89.1	496	5	AAU75490	Aau75490 S. aureus
11	2408	89.1	496	6	ADA89664	Ada89664 Staphyloc
12	2408	89.1	496	6	ADA89663	Ada89663 Staphyloc
13	1727	63.9	331	5	AAE22263	Aae22263 Staphyloc
14	1727	63.9	345	2	AAW31555	Aaw31555 Fibronect
15	520	19.2	961	6	ABU15854	Abu15854 Protein e
16	515	19.1	767	4	AAU34403	Aau34403 Staphyloc
17	515	19.1	767	4	AAU37547	Aau37547 Staphyloc
18	515	19.1	940	6	ABJ19111	Abj19111 Pathogen
19	515	19.1	948	6	ADA89470	Ada89470 Staphyloc
20	515	19.1	948	6	ABJ18923	Abj18923 Pathogen
21	515	19.1	948	6	ABM72536	Abm72536 Staphyloc
22	512	18.9	940	2	AAU07070	Aau07070 Fibronect
23	494.5	18.3	1166	2	AAU08643	Aau08643 S. aureus
24	494.5	18.3	1166	6	ABJ18982	Abj18982 Pathogen
25	494	18.3	1141	6	ABU42327	Abu42327 Protein e

26	472	17.5	1092	2	AAW41602	Aaw41602 Staphyloc
27	472	17.5	1092	7	ABM79019	Abm79019 Staphyloc
28	471.5	17.5	930	5	ABP40469	Abp40469 Staphyloc
29	471.5	17.5	931	3	AA583171	Aay83171 Cell wall
30	471.5	17.5	991	3	AAU70120	Aay70120 Staph. ep
31	468	17.3	1001	4	AAU37093	Aau37093 Staphyloc
32	467	17.3	567	6	ADA89555	Ada89555 Staphyloc
33	467	17.3	877	6	ADA89539	Ada89539 Staphyloc
34	467	17.3	877	6	ABU42504	Abu42504 Protein e
35	467	17.3	877	6	ABM72702	Abm72702 Staphyloc
36	467	17.3	913	6	ABJ18917	Abj18917 Pathogen
37	465.5	17.2	892	6	ABU42557	Abu42557 Protein e
38	465.5	17.2	892	8	ADO84849	Ado84849 S. epiderm
39	464	17.2	918	2	AAU08640	Aay08640 S. aureus
40	460.5	17.0	1018	4	AAU37245	Aau37245 Staphyloc
41	460.5	17.0	1018	4	AAU34301	Aau34301 Staphyloc
42	460.5	17.0	1018	6	ABJ18922	Abj18922 Pathogen
43	456	16.9	978	4	AAU33960	Aau33960 Staphyloc
44	455.5	16.9	549	7	ABM79020	Abm79020 Staphyloc
45	455.5	16.9	560	7	ABM79015	Abm79015 Staphyloc

ALIGNMENTS

RESULT 1
AAY58435
ID AAY58435 standard; protein; 933 AA.
XX AC AAY58435;
XX AC 27-MAR-2000 (first entry)
XX DT Staphylococcus aureus fibrinogen binding ClfA protein.
XX DE ClfA; fibrinogen binding protein; bacterial colonisation;
XX KW indwelling medical device; staphylococcal infection.
XX KW Staphylococcus aureus.
XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers
XX FT Peptide 1..39
XX FT Region 40..559 /note= "Signal peptide"
XX FT Region 332..550 /note= "Region A"
XX FT Region /note= "Fibrinogen-binding region"
XX FT Region 560..867 /note= "Region R"
XX FT Region 896..900 /note= "Gram positive wall-associated consensus motif"
XX US6008341-A.
XX 28-DEC-1999.
XX 22-AUG-1994; 94US-00293728.
XX 22-AUG-1994; 94US-00293728.
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX Foster TJ, Mcdevitt DL;
XX WPI; 2000-096389/08.
XX N-ESDB; AAZ55832.
XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus aureus, useful for treatment or prevention of infections.
XX Disclosure; Fig 2A-1-4; 35pp; English.
XX This sequence represents the Staphylococcus aureus fibrinogen-binding

appx 1.00

CC protein, ClfA. ClfA is an important receptor involved in S. aureus
 CC colonisation of indwelling medical devices (e.g., catheters, artificial
 CC heart valves). Shortly after implantation, the surfaces of medical
 CC devices become coated with host plasma and matrix proteins such as
 CC fibrinogen and fibronectin, and there is considerable evidence to suggest
 CC that bacterial adherence to fibrinogen/fibrin is important in the
 CC initiation of device-related infection. The fibrinogen-binding region of
 CC ClfA is thought to reside between residues 332 and 550 in a region R
 CC designated A. The protein also contains a repeated region (region R)
 CC comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus
 CC contains features present in surface proteins of other Gram positive
 CC bacteria that are responsible for anchoring the protein to the cell wall
 CC and cell membrane. ClfA, or its fragments, may be used to block S. aureus
 CC colonisation of wounds, to prevent adherence of S. aureus to indwelling
 CC medical devices, as vaccines to protect against S. aureus infection
 CC (e.g., mastitis in ruminants), to raise specific antibodies, and for
 CC diagnosis (by agglutination or immunoassay). The specific antibodies are
 CC used for passive immunisation, to block infection of wounds or adhesion
 CC of S. aureus and for diagnosis. Nucleotides encoding ClfA and its
 CC fragments may be used as diagnostic probes

XX Sequence 933 AA;

Query Match 100.0%; Score 2702; DB 3; Length 933;
 Best Local Similarity 100.0%; Pred. No. 3.7e-140;
 Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAPKTKDDTNVSDTKTSNT 60
 DB 23 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAPKTKDDTNVSDTKTSNT 82
 QY 61 NNGETSVAQNPAQQTQSSSTNATTEETPVTEATTTTNOANTPATQSSNTNAEELV 120
 DB 83 NNGETSVAQNPAQQTQSSSTNATTEETPVTEATTTTNOANTPATQSSNTNAEELV 142
 QY 121 NOTSETTFNDNTVSSVNSPQNSTNAENVTQDTSTATPSNNEAPQSTDAKNQV 180
 DB 143 NOTSETTFNDNTVSSVNSPQNSTNAENVTQDTSTATPSNNEAPQSTDAKNQV 202
 QY 181 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTVPHQAGYKLVNG 240
 DB 203 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTVPHQAGYKLVNG 262
 QY 241 FSVPSNAVKGDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 300
 DB 263 FSVPSNAVKGDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 322
 QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLVDYKYGKFPYNSIKG 360
 DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLVDYKYGKFPYNSIKG 382
 QY 361 TTDQIDKTNNTYRQTIYVNPNGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKVDNAA 420
 DB 383 TTDQIDKTNNTYRQTIYVNPNGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKVDNAA 442
 QY 421 DLSESYFVNPENFEDVTSNVTITFPNPNQYKVEFTPDQITTPYTVVNGHIDPNSKGD 480
 DB 443 DLSESYFVNPENFEDVTSNVTITFPNPNQYKVEFTPDQITTPYTVVNGHIDPNSKGD 502
 QY 481 LAIRSTLYGNSNIIWRSMNSWNEVAFNNGSGSGDGIDKPVVPEQDPE 528
 DB 503 LAIRSTLYGNSNIIWRSMNSWNEVAFNNGSGSGDGIDKPVVPEQDPE 550

RESULT 2

AAB69508
 ID AAB69508 standard; protein; 933 AA.
 XX
 AC AAB69508;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Staphylococcus aureus ClfA protein.

XX Staphylococcus aureus; clfA; antibiotic; vaccine;
 KW fibrinogen binding protein; bacterial infection; mastitis.
 XX
 OS Staphylococcus aureus.
 XX
 PN US6177084-B1.
 XX
 PD 23-JAN-2001.
 XX
 PF 19-OCT-1999; 99US-00421869.
 XX
 PR 22-AUG-1994; 94US-00293728.
 XX
 PA (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX
 PI Foster TJ, Medevitt DL;
 XX
 XX WPI; 2001-181608/18.
 DR N-PSDB; AAF58593.
 XX
 PT Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
 to prevent infection, promote wound healing, block adherence to
 PT indwelling medical devices and for diagnosing staphylococcus aureus
 PT infection.
 XX
 PS Claim 5; Fig 2; 30pp; English.
 XX
 CC The present sequence is a novel Staphylococcus aureus fibrinogen binding
 CC protein. It is useful as a vaccine to protect against human and animal
 CC infections caused by S.aureus, such as against mastitis, to block
 CC S.aureus from colonising and infecting a wound, to block adherence of
 CC S.aureus to indwelling medical devices such as catheters, replacement
 CC heart valves and cardiac assist devices. The protein can be used to
 CC diagnose bacterial infections. Polyclonal and monoclonal antibodies
 CC raised against the fibrinogen binding protein are useful for passively
 CC immunising against infections caused by S.aureus, to prevent infection of
 CC a wound and to diagnose bacterial infections
 XX
 SQ Sequence 933 AA;

Query Match 100.0%; Score 2702; DB 4; Length 933;
 Best Local Similarity 100.0%; Pred. No. 3.7e-140;
 Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAPKTKDDTNVSDTKTSNT 60
 DB 23 VGTLLGFGLLSKKEADASENSVTQSDASNESKSDSSVSAPKTKDDTNVSDTKTSNT 82
 QY 61 NNGETSVAQNPAQQTQSSSTNATTEETPVTEATTTTNOANTPATQSSNTNAEELV 120
 DB 83 NNGETSVAQNPAQQTQSSSTNATTEETPVTEATTTTNOANTPATQSSNTNAEELV 142
 QY 121 NOTSETTFNDNTVSSVNSPQNSTNAENVTQDTSTATPSNNEAPQSTDAKNQV 180
 DB 143 NOTSETTFNDNTVSSVNSPQNSTNAENVTQDTSTATPSNNEAPQSTDAKNQV 202
 QY 181 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTVPHQAGYKLVNG 240
 DB 203 NOAVNTSAPRMRAFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTVPHQAGYKLVNG 262
 QY 241 FSVPSNAVKGDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 300
 DB 263 FSVPSNAVKGDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 322
 QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLVDYKYGKFPYNSIKG 360
 DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGSTTANKVLVDYKYGKFPYNSIKG 382
 QY 361 TTDQIDKTNNTYRQTIYVNPNGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKVDNAA 420
 DB 383 TTDQIDKTNNTYRQTIYVNPNGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKVDNAA 442

Appendix

QY 421 DLSESYFVNPFENFEDVTNSNIIFPNPNQYKVEFTPDQOITTPYIVVNGHIDPNSKGD 480
DB 443 DLSESYFVNPFENFEDVTNSNIIFPNPNQYKVEFTPDQOITTPYIVVNGHIDPNSKGD 502
QY 481 LALRSTLYGYNISIIWRSMDSWDNEVAFNNGSGDGDIDKFPVPEQDPE 528
DB 503 LALRSTLYGYNISIIWRSMDSWDNEVAFNNGSGDGDIDKFPVPEQDPE 550

RESULT 3
ABJ18947
ID ABJ18947 standard; protein; 933 AA.
XX
XX ABJ18947;
XX
XX 06-MAR-2003 (first entry)
XX
XX Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
XX
XX Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
XX hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
XX auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
XX autoimmune disease; HIV; hepatitis.
XX
XX Staphylococcus sp.
XX
XX WO200259148-A2.
XX
XX 01-AUG-2002.
XX
XX 21-JAN-2002; 2002WO-EP000546.
XX
XX 26-JAN-2001; 2001AT-00000130.
XX
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
XX Meinke A, Nagy E, Von Ahseu W, Klade C, Henics T, Zauner W;
XX Mink DB, Vytvyska O, Etz W, Dryla A, Weichhart T, Hafner M;
XX Tempelmeier B;
XX
XX WPI; 2003-075410/07.
XX
XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.
XX
XX Example 7; Page 160; 252pp; English.
XX
XX The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX allergen, a tissue or host prone to auto-immunity, where the antigens are
XX used in a vaccine, comprises providing antibody preparation from a plasma
XX pool of a type of animal, or individual sera with antibodies against the
XX specific pathogen, tumour, allergen, tissue or host prone to auto-
XX immunity. The hyperimmune serum-reactive antigens comprising any of the
XX 62 sequences of 53-2261 amino acids fully defined in the specification,
XX or their hyperimmune fragments are useful for the manufacture of a
XX pharmaceutical preparation, particularly a vaccine against staphylococcal
XX infections or colonisation against S. aureus or S. epidermidis. The
XX preparation of antibodies is useful for the manufacture of a medicament
XX for treating or preventing staphylococcal infections or colonisation
XX against S. aureus or S. epidermidis. The antibody preparations may also
XX be used for diagnostic and imaging purposes. Other conditions that can be
XX treated include cancer, autoimmune diseases or infections caused by viral
XX (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX sequence represents a staphylococcal protein relating to the method for
XX identifying and producing pathogen specific antigens of the invention

Query Match 99.7%; Score 2694; DB 6; Length 933;
Best Local Similarity 99.8%; Pred. No. 1e-139;

Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VGLTIGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 60
DB 23 VGLTIGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETTQSSSTNATTEETPVTGEATTTTINQANTPATTOSSTNAEELV 120
DB 83 NNGETSAQNPAQOETTQSSSTNATTEETPVTGEATTTTINQANTPATTOSSTNAEELV 142
QY 121 NQTSNETTNDTNTVSVNSPQNSTNAENVSTTQDTSSTATPSNNSAPOSTDASNKDVV 180
DB 143 NQTSNETTNDTNTVSVNSPQNSTNAENVSTTQDTSSTATPSNNSAPOSTDASNKDVV 202
QY 181 NQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHAGYVKLNLYG 240
DB 203 NQAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHAGYVKLNLYG 262
QY 241 FSVPSNAVGKDTFKITVPKELNLNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYTFDY 300
DB 263 FSVPSNAVGKDTFKITVPKELNLNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYTFDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKFKFNLSIKG 360
DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKFKFNLSIKG 382
QY 361 TIDQIDKNTNTYRQTIYVNPFGDNVIAVLTKNLPNTDSNALIDQOQNTSIVKVKVDNAA 420
DB 383 TIDQIDKNTNTYRQTIYVNPFGDNVIAVLTKNLPNTDSNALIDQOQNTSIVKVKVDNAA 442
QY 421 DLSESYFVNPFENFEDVTNSNIIFPNPNQYKVEFTPDQOITTPYIVVNGHIDPNSKGD 480
DB 443 DLSESYFVNPFENFEDVTNSNIIFPNPNQYKVEFTPDQOITTPYIVVNGHIDPNSKGD 502
QY 481 LALRSTLYGYNISIIWRSMDSWDNEVAFNNGSGDGDIDKFPVPEQDPE 528
DB 503 LALRSTLYGYNISIIWRSMDSWDNEVAFNNGSGDGDIDKFPVPEQDPE 550

RESULT 4
ABM72221
ID ABM72221 standard; protein; 927 AA.
XX
XX AC ABM72221;
XX
XX DT 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus protein #1461.
XX
XX DE Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX KW enzymatic assay; antibiotic target.
XX
XX OS Staphylococcus aureus.
XX
XX PN WO200294868-A2.
XX
XX PD 28-NOV-2002.
XX
XX PF 27-MAR-2002; 2002WO-IB002637.
XX
XX PR 27-MAR-2001; 2001GB-00007661.
XX
XX PA (CHIR-) CHIRON SPA.
XX
XX PI Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120785/11.
XX
XX DR N-PSDB; ACF73781.
XX
XX PT New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.

PS Claim 1; SEQ ID NO 2922; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX
SQ Sequence 927 AA;

Query Match 99.6%; Score 2690; DB 6; Length 927;

Best Local Similarity 99.6%; Pred. No. 1.7e-139;

Matches 526; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAPKTDITNVSDTKTSNT 60
DB 23 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAPKTDITNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 120
DB 83 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 142
QY 121 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNEAPQSDASNKQVY 180
DB 143 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNEAPQSDASNKQVY 202
QY 181 NQAVNTSAPRMRAPFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 240
DB 203 NQAVNTSAPRMRAPFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 262
QY 241 FSVNPSAVKGDFTFKITVPKELNMGVTSKAPVPPIMAGDQVLANGVSDGNVITFTDY 300
DB 263 FSVNPSAVKGDFTFKITVPKELNMGVTSKAPVPPIMAGDQVLANGVSDGNVITFTDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGISTTANKTVLDVDEYKGYFYNLSIKG 360
DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGISTTANKTVLDVDEYKGYFYNLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAA 420
DB 383 TIDQIDKTNNTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAA 442
QY 421 DLSESYFVNPFEDVNTSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 480
DB 443 DLSESYFVNPFEDVNTSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNSKGD 502
QY 481 LALRSTLYGNSNIIWRSMWDEVAFNNGSGSGDIDKPVVPEQFDE 528
DB 503 LALRSTLYGNSNIIWRSMWDEVAFNNGSGSGDIDKPVVPEQFDE 550

RESULT 5

AAW89801

ID AAW89801 standard; protein; 936 AA.

XX

AC AAW89801;

DT

XX 16-MAR-1999 (first entry)

DE

XX Staphylococcus aureus protein SEQ ID #5249.

XX

KW Computer readable medium; vaccine; S. aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome.

XX

OS Staphylococcus aureus.

XX

PN EP786519-A2.

XX

PD 30-JUL-1997.

XX

PF 07-JAN-1997; 97EP-00100117.

XX

PR 05-JAN-1996; 96US-0009861P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX

WPI; 1997-374922/35.

XX

PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of anti-
PT S. aureus vaccines.

XX

PS Claim 23; Page 3255-3258; 3271pp; English.

XX

CC This sequence represents a Staphylococcus aureus protein sequence of the
CC invention. The DNA sequences encoding the S. aureus proteins are recorded
CC on a computer readable medium, preferably selected from a floppy or hard
CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
CC Homology searches using the S. aureus DNA sequences allows putative
CC functions to be assigned so that protein-encoding or regulatory regions
CC of commercial, therapeutic or industrial importance can be obtained.
CC Specifically, sequences which are likely to encode antigens have been
CC identified and these polypeptides can be used in a vaccine composition
CC against S. aureus infection. The polypeptides can also be used in a kit
CC for the immunodetection of S. aureus in a sample. S. aureus is implicated
CC in numerous human diseases, including cellulitis, eyelid infections, food
CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
CC DNA sequences can be used for recombinant production of the polypeptides.
CC The new DNA sequences (and their fragments) are useful as primers or
CC probes for isolating homologues of any of the 5191 S. aureus DNA sequences
CC contained on the computer readable medium

SQ Sequence 936 AA;

Query Match

Best Local Similarity 99.6%; Score 2690; DB 2; Length 936;

Matches 526; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAPKTDITNVSDTKTSNT 60
DB 32 VGTLLGFLSSKEADASENSVTQSDASNESKNDSSVSAPKTDITNVSDTKTSNT 91
QY 61 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 120
DB 92 NNGETSAQNPAQOQTQSSSTNATTEPTVTGEATTTTNOANTPATQSSNTNAEELV 151
QY 121 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNEAPQSDASNKQVY 180
DB 152 NOTSETTNDNTVSSVNSPQNSINAENVSTQDTSATPSNNEAPQSDASNKQVY 211
QY 181 NQAVNTSAPRMRAPFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 240
DB 212 NQAVNTSAPRMRAPFSLAAVAADAPAGTDTITNQLTNVTGIDSGTTPVPHQAGYKLVNG 271
QY 241 FSVNPSAVKGDFTFKITVPKELNMGVTSKAPVPPIMAGDQVLANGVSDGNVITFTDY 300
DB 272 FSVNPSAVKGDFTFKITVPKELNMGVTSKAPVPPIMAGDQVLANGVSDGNVITFTDY 331
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGISTTANKTVLDVDEYKGYFYNLSIKG 360
DB 332 VNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGISTTANKTVLDVDEYKGYFYNLSIKG 391
QY 361 TIDQIDKTNNTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAA 420
DB 392 TIDQIDKTNNTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAA 451

CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 935 AA;

Query Match 91.3%; Score 2466; DB 6; Length 935;
Best Local Similarity 91.1%; Pred. No. 3 4e-127;
Matches 481; Conservative 19; Mismatches 28; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKEADASENSVTQSDASNEKSDSSVSAAKPTDDTNVSDTKTSNT 60
DB 23 VGTLLGFGLLSKEADASENSVTQSDASNEKSDSSVSAAKPTDDTNVSDTKTSNT 82
QY 61 NNGETSVQAONPAQOETQSSSTNATEETPVGTGEATTTTNCANTPATTQSSNTARELV 120
DB 83 NNGETSVQAONPAQOETQSSSTNATEETPVGTGEATTTTNCANTPATTQSSNTARELV 142
QY 121 NQTSNETTNDNTVSSVNSPQNSNAENSVTTQDTSTTEATPSNESAQSDASNKQVV 180
DB 143 NQTSNETTNDNTVSSVNSPQNSNAENSVTTQDTSTTEATPSNESAQSDASNKQVV 202
QY 181 NOAVNTSAPRMRAFSLAAVAADAPAGTDITNQLNNTVVGIDSGTIVPHQAGYKLVNG 240
DB 203 SOAVNPSTPRMRAFSLAAVAADAPAGTDITNQLTDVKVITDSTGTTVPHQAGYKLVNG 262
QY 241 FSVPSAVKGDFTKIVTKPELNLNGVTSTAKPPIMAGDOVLANGVIDSDGNVITFTDY 300
DB 263 FSVPSAVKGDFTKIVTKPELNLNGVTSTAKPPIMAGDOVLANGVIDSDGNVITFTDY 322
QY 301 VNTKDDVKATLTPAYIDPENVKKTGNVTLATIGISTTANKVLVDYERKGFYKLSIKG 360
DB 323 VDKNENVTANITMPAYIDPENVKKTGNVTLATIGISTTANKVLVDYERKGFYKLSIKG 382
QY 361 TTQIDKTNNTVROTIIYNPNSGDNVIAVLGTNKLKPTDSDNALIDQOQNTSIKVKYVDNAA 420
DB 383 TTQIDKTNNTVROTIIYNPNSGDNVIAVLGTNKLKPTDSDNALIDQOQNTSIKVKYVDNAN 442
QY 421 DLSESYFVNPENFEDVTSVNTIFPNQYKVEENTPDDOITTPYIVVNGHIDNSKGD 480
DB 443 DLSESYFVNPENFEDVTSVNTIFPNQYKVEENTPDDOITTPYIVVNGHIDPASTGD 502
QY 481 LALRSTLYGNGNIIWRMSWNEVAFNNGSGSGDIDKPVVPEQDPE 528
DB 503 LALRSTFYGYDGNFIWRMSWNEVAFNNGSGSGDIDKPVVPEQDPE 550

RESULT 8

AAU33975
ID AAU33975 standard; protein; 1021 AA.
XX
AC AAU33975;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #251.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
OS Staphylococcus aureus.
PN WO200170955-A2.
XX 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AASS1834.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 5471; 511pp; English.
XX

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 1021 AA;

Query Match 90.7%; Score 2450; DB 4; Length 1021;
Best Local Similarity 90.3%; Pred. No. 2 9e-126;
Matches 477; Conservative 22; Mismatches 29; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKEADASENSVTQSDASNEKSDSSVSAAKPTDDTNVSDTKTSNT 60
DB 23 VGTLLGFGLLSKEADASENSVTQSDASNEKSDSSVSAAKPTDDTNVSDTKTSNT 82
QY 61 NNGETSVQAONPAQOETQSSSTNATEETPVGTGEATTTTNCANTPATTQSSNTARELV 120

Db 83 NNGETSAQNPAQOQETTSASNTATTEETPTVGTATTTATKQANTPATTQSSNTNABEVL 142
QY 121 NQTSNTEFTNDTNTVSSVNSPQNSTNAENVSITQDTSSTATPSNNSAPOSTDASNKDVV 180
Db 143 NQTSNETASNDTNTVSSVNSPQNSTNAENVSITQDTSSTATPSNNSAPOSTDASNKDVV 202
QY 181 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYPHQAQGVKLNLYG 240
Db 203 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYPHQAQGVKLNLYG 262
QY 241 FSPVNSAVKGDGTFKITVPKELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 300
Db 263 FSPVNSAVKGDGTFKITVPKELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKGTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKG 360
Db 323 VDTKENVTANITMPAYIDPENVKGTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQTSIKVYKVDNAA 420
Db 383 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQTSIKVYKVDNAN 442
QY 421 DLSSEYFVNPENFEDVNTSNVITFPNPNQYKVFNTDQDQITTPYIVVNGHIDPNKSGD 480
Db 443 DLSSEYVNPSPDFEDVTNQVRISFPNANQYKVFPTDDQITTPYIVVNGHIDPASTGD 502
QY 481 LALRSTLYGNSNLIWRSMWDEVAFNNGSGGSGDGIDKVPVPEQDPE 528
Db 503 LALRSTLYGNSNLIWRSMWDEVAFNNGSGGSGDGIDKVPVPEQDPE 550

RESULT 9
AAU36951
ID AAU36951 standard; protein; 1021 AA.
XX
AC AAU36951;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1121.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Heselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS54810.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 12544; 511pp; English.
XX

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1021 AA;
Query Match 90.7%; Score 2450; DB 4; Length 1021;
Best Local Similarity 90.3%; Pred. No. 2.9e-126;
Matches 477; Conservative 22; Mismatches 29; Indels 0; Gaps 0;
QY 1 VGTLLIGFLLSSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 60
Db 23 VGTLLIGFLLSSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOQETTSASNTATTEETPTVGTATTTATKQANTPATTQSSNTNABEVL 120
Db 83 NNGETSAQNPAQOQETTSASNTATTEETPTVGTATTTATKQANTPATTQSSNTNABEVL 142
QY 121 NQTSNTEFTNDTNTVSSVNSPQNSTNAENVSITQDTSSTATPSNNSAPOSTDASNKDVV 180
Db 143 NQTSNETASNDTNTVSSVNSPQNSTNAENVSITQDTSSTATPSNNSAPOSTDASNKDVV 202
QY 181 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYPHQAQGVKLNLYG 240
Db 203 NQAVNTSAPMRAPFSLAAVAADAPAAAGTDTITNLTNTVTVGIDSGTTVYPHQAQGVKLNLYG 262
QY 241 FSPVNSAVKGDGTFKITVPKELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 300
Db 263 FSPVNSAVKGDGTFKITVPKELNLNGVTSTAKVPPIMAGDOVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKGTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKG 360
Db 323 VDTKENVTANITMPAYIDPENVKGTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQTSIKVYKVDNAA 420
Db 383 TIDQIDKTNNTYRQTIYVNPVSGDNVIAPIVLTGNLKPNTDSNALIDQOQTSIKVYKVDNAN 442
QY 421 DLSSEYFVNPENFEDVNTSNVITFPNPNQYKVFNTDQDQITTPYIVVNGHIDPNKSGD 480
Db 443 DLSSEYVNPSPDFEDVTNQVRISFPNANQYKVFPTDDQITTPYIVVNGHIDPASTGD 502
QY 481 LALRSTLYGNSNLIWRSMWDEVAFNNGSGGSGDGIDKVPVPEQDPE 528
Db 503 LALRSTLYGNSNLIWRSMWDEVAFNNGSGGSGDGIDKVPVPEQDPE 550

RESULT 10
AAU75490
ID AAU75490 standard; protein; 496 AA.
XX
AC AAU75490;
XX
DT 23-APR-2002 (first entry)
XX
DE S. aureus antigenic protein associated protein #10.
XX
KW Antigenic protein; vaccine; SEREX; antibacterial; antiinflammatory;

dermatological; antiulcer; tuberculostatic; immunosuppressive;
septicemia; food poisoning; skin disorders; peritonitis; endocarditis;
tuberculosis; blood infection; sepsis; meningitis; pneumonia;
stomach ulcer; gonorrhea; necrotising fasciitis; impetigo;
Lyme's disease; gastro-enteritis; dysentery; shigellosis.
Staphylococcus aureus.
WO2001198499-A1.
27-DEC-2001.
20-JUN-2001; 2001WO-GB002685.
20-JUN-2000; 2000GB-00014907.
(UYSH-) UNIV SHEPFIELD.
(BIOS-) BIOSYNEXUS INC.
Foster S, McDowell P, Brummell K, Clarke S;
WPI; 2002-106544/14.
Identifying antigenic polypeptides expressed by pathogenic organisms
e.g., Staphylococcus aureus during infection, by SEREX (serological
identification of antigens by recombinant expression cloning) techniques.
Disclosure; Page 72-74; 85pp; English.
The invention relates to a method for identifying antigenic polypeptides
expressed by pathogenic organisms e.g., Staphylococcus aureus during
infection, by SEREX (serological identification of antigens by
recombinant expression cloning) techniques. The method involves providing
a nucleic acid library encoding genes/partial gene sequences of
pathogenic organisms, transforming/transfecting the library into host
cells, contacting the polypeptides expressed by the genes with autologous
antisera derived from an animal infected with, or has been infected with
the pathogen and purifying the nucleic acid encoding the polypeptide or
partial polypeptide binding to the antisera. Also included are the
nucleic acids and polypeptides isolated by the method, vectors and
transformed cells expressing them, a vaccine comprising the polypeptide,
and the production of monoclonal antibodies against the polypeptides. The
protein and vaccine are useful for immunising an animal (preferably
human) against a pathogenic microbe. The proteins and antibodies are
useful for manufacturing a medicament for treating Staphylococcus aureus-
associated septicemia, food poisoning or skin disorders or
Staphylococcus epidermidis-associated septicemia, peritonitis,
endocarditis, tuberculosis, blood infections, sepsis, meningitis,
pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
sequence is an S. aureus protein sequence associated with the antigenic
proteins of the invention. Note: The present sequence is included in the
sequence listing but is not mentioned anywhere else in the specification
Sequence 496 AA;
Query Match 89.1%; Score 2408; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.4e-124;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 VGTLLGFLGSKKEADASENSVTQSDASNSKNSDSSVSAAPKTDNTNSDTSSNT 60
23 VGTLLGFLGSKKEADASENSVTQSDASNSKNSDSSVSAAPKTDNTNSDTSSNT 82
61 NNGETSVAQNPAQQTQSSSTNAETPTVGEATTTTQANTPATQSSNTNAEELV 120
83 NNGETSVAQNPAQQTQSSSTNAETPTVGEATTTTQANTPATQSSNTNAEELV 142
121 NQTSNETTFNDNTVSSVNSPONSNAENSVTQDTSATPSPNNESAPQSDASNKDQV 180
143 NQTSNETTFNDNTVSSVNSPONSNAENSVTQDTSATPSPNNESAPQSDASNKDQV 202
181 NQAVNTSAPRMEAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTVPYHQAGYVKLNYG 240

Db 203 NQAVNTSAPRMEAFSLAAVAADAPAGTDITNQLTNVTVGIDSGTTVPYHQAGYVKLNYG 262
Qy 241 FSVPN SAVKGDFTFKITVPKELNNGVTSTAKVPPTIMAGDQVLANGVIDSDGNVITFTDY 300
Db 263 FSVPN SAVKGDFTFKITVPKELNNGVTSTAKVPPTIMAGDQVLANGVIDSDGNVITFTDY 322
Qy 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTGLVDYKYKGFYNLSIKG 360
Db 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTGLVDYKYKGFYNLSIKG 382
Qy 361 TIDQIDKNTTYRQTIYVNPNGDNVIAVLTKNLKPTDSNALIDQONTSIKYKVDNAA 420
Db 383 TIDQIDKNTTYRQTIYVNPNGDNVIAVLTKNLKPTDSNALIDQONTSIKYKVDNAA 442
Qy 421 DLSESYFVNPFEDVTNSVNIITFPNPNQYKVFNTDDOITTPYIVVNVGHID 474
Db 443 DLSESYFVNPFEDVTNSVNIITFPNPNQYKVFNTDDOITTPYIVVNVGHID 496
RESULT 11
ADA89664
ID ADA89664 standard; protein; 496 AA.
XX
ADA89664;
XX
20-NOV-2003 (first entry)
XX
XX
Staphylococcus aureus antigenic protein #203.
XX
antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antitumor; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteremia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.
XX
Staphylococcus aureus.
OS
WO2003011899-A2.
PN
13-FEB-2003.
XX
02-AUG-2002; 2002WO-GB003606.
XX
02-AUG-2001; 2001GB-00018825.
PR
09-JAN-2002; 2002GB-00000349.
XX
(UYSH-) UNIV SHEPFIELD.
PA
(BIOS-) BIOSYNEXUS INC.
XX
Foster S, Mond J, Clarke S, McDowell P, Brummell K;
PI
WPI; 2003-256434/25.
XX
New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
DR useful as a vaccine for immunizing humans against e.g. bacteremia, septic
XX shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or
PT impetigo.
XX
Claim 4; Page 161; 189pp; English.
PS
The present invention describes an antigenic protein or its part, which
XX is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
CC sequences (designated dnaA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;

CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, immunoprotective, immunosuppressive,
CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
CC bacteria-associated food poisoning, blood infections, peritonitis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicaemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus antigenic
CC protein sequence from the present invention.
XX
SQ Sequence 496 AA;

Query Match 89.1%; Score 2408; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.4e-124;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VTGLIGFGLSSKEADASENSVTQDSASNESKNSDSSVSAAKPTDDTNVSDTKTSNT 60
DB 23 VTGLIGFGLSSKEADASENSVTQDSASNESKNSDSSVSAAKPTDDTNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETTQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 120
DB 83 NNGETSAQNPAQOETTQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 142
QY 121 NQTSNETTFNDTNVSSVNSPQNSTNAENVSTTQDTSTEATPPSNESAPQSTDAENKDVV 180
DB 143 NQTSNETTFNDTNVSSVNSPQNSTNAENVSTTQDTSTEATPPSNESAPQSTDAENKDVV 202
QY 181 NQAVNTSAPMRAPFSLAAVAADAPAACTDTTNQLTNTVTVGIDSGTTVYPHQAGVVKLYNG 240
DB 203 NQAVNTSAPMRAPFSLAAVAADAPAACTDTTNQLTNTVTVGIDSGTTVYPHQAGVVKLYNG 262
QY 241 FSPVNSAVKGDFTKITVPKELNGLNGVSTAKVPPIMAGDQVLANGVIDSGNVLVYTFDY 300
DB 263 FSPVNSAVKGDFTKITVPKELNGLNGVSTAKVPPIMAGDQVLANGVIDSGNVLVYTFDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDEYKYGKFNLSIKG 360
DB 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDEYKYGKFNLSIKG 382
QY 361 TIQIDKTNNTYRTIIVNPSGDNVIAPLVTGNLKPNTDSNALIDQNTSIKVKYVDNAA 420
DB 383 TIQIDKTNNTYRTIIVNPSGDNVIAPLVTGNLKPNTDSNALIDQNTSIKVKYVDNAA 442
QY 421 DLSESFYVNPENFEDVNSVNIITPPNQYKVEFNTDDQITTYIIVVNGHID 474
DB 443 DLSESFYVNPENFEDVNSVNIITPPNQYKVEFNTDDQITTYIIVVNGHID 496

RESULT 12

ADA89663

ID ADA89663 standard; protein; 496 AA.

XX

AC ADA89663;

XX

DT 20-NOV-2003 (first entry)

XX

DE Staphylococcus aureus antigenic protein #202.

XX

KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;

KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;

KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;

KW bacteraemia; septic shock; organ infection; skin infection;

KW bacterial basal colonisation; bacterial eye infection; septicaemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.
OS Staphylococcus aureus.
XX
PN W02003011899-A2.
XX
PD 13-FEB-2003.
XX
PF
XX 02-AUG-2002; 2002WO-GH003606.
XX
PR 02-AUG-2001; 2001GB-00018825.
PR 09-JAN-2002; 2002GB-00000349.
XX
PA (UYSH-) UNIV SHEFFIELD.
PA (BIOS-) BIOSYNEXUS INC.
XX
XX Foster S, Mond J, Clarke S, McDowell P, Brummel K;
XX
XX WPI; 2003-256434/25.
XX
PT New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
PT useful as a vaccine for immunizing humans against e.g. bacteraemia, septic
PT shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
PT impetigo.
XX
PS Claim 4; Page 160-161; 189pp; English.

The present invention describes an antigenic protein or its part, which
is for use as a vaccine. The antigenic protein is encoded by an isolated
DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
sequences (designated dnaSA and dna SE, respectively; and which encodes a
protein expressed by a pathogenic organism. Also described: (1) a vaccine
composition comprising at least one antigenic protein; (2) a method of
immunising an animal against a disease or condition caused by a
pathogenic microbe by administering the antigenic protein or the vaccine;
(3) an antibody or its binding part obtainable by the method above; (4)
preparing a hybridoma cell line producing monoclonal antibodies; (5) a
hybridoma cell line produced by the method of (4); and (6) identifying
opsonic antigens expressed by a pathogenic microbe. The antigenic
proteins have antibacterial, neuroprotective, immunosuppressive,
antiinflammatory, antiulcer, immunostimulant and ophthalmological
activities, and can be used in vaccines. The antigenic proteins or
vaccines can be used for immunising an animal (specifically a human)
against a disease or condition caused by a pathogenic microbe, e.g.
bacteraemia, septic shock, organ infection, skin infection, bacterial
basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
bacteria-associated food poisoning, blood infections, peritonitis,
endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
strep throat, streptococcal-associated toxic shock, necrotising
fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
skin disorders, S. epidermidis-associated septicaemia, peritonitis or
endocarditis. The present sequence represents a S. aureus antigenic
protein sequence from the present invention.

Sequence 496 AA;

Query Match

Best Local Similarity 89.1%; Score 2408; DB 6; Length 496;

Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTGLIGFGLSSKEADASENSVTQDSASNESKNSDSSVSAAKPTDDTNVSDTKTSNT 60
DB 23 VTGLIGFGLSSKEADASENSVTQDSASNESKNSDSSVSAAKPTDDTNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETTQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 120
DB 83 NNGETSAQNPAQOETTQSSSTNATTEETPTVGEATTTTNTQANTPATTOSSTNABELV 142

QY 121 NOTSETTENDTNTVSSVNSPQNSTAENSVTQDTSTTEATPSNNEAPQSDASNKQV 180
Db 143 NOTSETTENDTNTVSSVNSPQNSTAENSVTQDTSTTEATPSNNEAPQSDASNKQV 202
QY 181 NOAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLTNVTVGIDSGTTVPYHQAGYKLYNG 240
Db 203 NOAVNTSAPRMRAFSLAAVAADAPAAAGTDITNQLTNVTVGIDSGTTVPYHQAGYKLYNG 262
QY 241 RSPVNSAVKGDFTKIVPKELNGLVGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 300
Db 263 RSPVNSAVKGDFTKIVPKELNGLVGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDY 322
QY 301 VNTKDDVKATLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKG 360
Db 323 VNTKDDVKATLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKG 382
QY 361 TIDQIDKTNNTVROTIYVNPNGDNVIAVLTGNLKPENTDSNALIDQONTSIKVKVDNAA 420
Db 383 TIDQIDKTNNTVROTIYVNPNGDNVIAVLTGNLKPENTDSNALIDQONTSIKVKVDNAA 442
QY 421 DLSSEYFVNPENFEDVTVSNITFPNPKYKVEFNTDDQITTPYVNVNGHID 474
Db 443 DLSSEYFVNPENFEDVTVSNITFPNPKYKVEFNTDDQITTPYVNVNGHID 496

RESULT 13
AAE29263
ID AAE29263 standard; protein; 331 AA.
XX
AC AAE29263;

DT 27-JAN-2003 (first entry)

Staphylococcus aureus Clf33 protein.

Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein;
immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

Staphylococcus aureus.

WO200272600-A2.

19-SEP-2002.

28-JAN-2002; 2002WO-US002296.

26-JAN-2001; 2001US-0264072P.

12-MAR-2001; 2001US-0274611P.

18-JUN-2001; 2001US-0298413P.

30-JUL-2001; 2001US-0308116P.

(INH-) INHIBITEX INC.

Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;

WPI; 2002-759834/82.

N-PSDB; AAD46862.

New anti-clumping factor A (ClfA) monoclonal antibody, useful for

treating or preventing Staphylococcus aureus infection e.g. wound

infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

a human or animal.

Claim 9; Page 70-72; 80pp; English.

The invention relates to monoclonal antibody which binds the clumping

factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA

monoclonal antibody is useful for treating or preventing S. aureus

infection in a human or animal, and for inhibiting the binding of

staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment

S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3

protein is useful for inducing an immunological response in a human or

CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is Staphylococcus aureus Clf33 protein
XX
SQ Sequence 331 AA;

Query Match 63.9%; Score 1727; DB 5; Length 331;
Best Local Similarity 100.0%; Pred. No. 3.9e-87;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 VAADAPAAAGTDITNQLTNVTVGIDSGTTVPYHQAGYKLYNGSVNPENFEDVTN 258

Db 2 VAADAPAAAGTDITNQLTNVTVGIDSGTTVPYHQAGYKLYNGSVNPENFEDVTN 61

QY 259 KELNGLVGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDVNTKDDVKATLTPAYID 318

Db 62 KELNGLVGTSTAKVPPIMAGDQVLANGVIDSDGNVITFTDVNTKDDVKATLTPAYID 121

QY 319 PENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYQTIV 378

Db 122 PENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYQTIV 181

QY 379 NPNGDNVIAVLTGNLKPENTDSNALIDQONTSIKVKVDNAAADLSEYFVNPENFEDVTN 438

Db 182 NPNGDNVIAVLTGNLKPENTDSNALIDQONTSIKVKVDNAAADLSEYFVNPENFEDVTN 241

QY 439 SVNITFPNPKYKVEFNTDDQITTPYVNVNGHIDPNSKGLALRSTLYGYNLSIWR 498

Db 242 SVNITFPNPKYKVEFNTDDQITTPYVNVNGHIDPNSKGLALRSTLYGYNLSIWR 301

QY 499 MSWNEVAFNNGSGSGDGIDKPVVPEQDE 528

Db 302 MSWNEVAFNNGSGSGDGIDKPVVPEQDE 331

RESULT 14

AAW31555

ID AAW31555 standard; protein; 345 AA.

XX AAW31555;

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 21-MAY-1998 (first entry)

DE Fibronectin-binding MSCRAMM derivative pCF33.

XX Fibronectin; pCF33; collagen binding protein; sepsis; infection;

XX microbial surface component regionising adhesive matrix molecule; MSCRAMM;

XX adhesin; vaccine; immunisation; diagnosis; therapy; mastitis.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

FT Peptide 1..12

FT /note= "vector pOE30-derived peptide"

XX WO9743314-A2.

PN 20-NOV-1997.

XX 14-MAY-1997; 97WO-US008210.

XX 16-MAY-1996; 96US-0017678P.

XX (TEXA-) UNIV TEXAS A & M SYSTEM.

PA (UABR-) UAB RES FOUND.

XX Hoeek M, Patti JM, House-Pompeo K, Sthanam N, Symersky J;

XX WPI; 1998-008801/01.

XX Antibody that interacts with collagen binding domain of Staphylococcal

PT cna gene product - useful to prevent bacterial sepsis in animal infected
PT with *Staphylococcus aureus*.
XX
PS Disclosure; Page 91; 143pp; English.
XX

CC This protein comprises *Staphylococcus aureus* fibronectin-binding
CC microbial surface component recognising adhesive matrix molecule (MSCRAMM)
CC derivative pCF33, plus a vector-derived N-terminal peptide. The invention
CC relates to claimed nucleic acid sequences (see AAT93436-38) encoding S.
CC aureus collagen binding protein (CBP) epitopes M17, M31 and M55 (see
CC AAW31552-54) that confer protection against *S. aureus* infection. CBP
CC protein and antigenic epitopes are contemplated for use in the treatment
CC of pathological infections, especially to prevent bacterial adhesion to
CC collagen. The epitopes are also contemplated for use in the preparation
CC of vaccines and as carrier proteins in vaccine formulations, as well as
CC in the formulation of compositions for the prevention of *S. aureus*
CC infection. pCF33 and pQD (see AAW31556) were used to raise anti-MSCRAMM
CC polyclonal antibodies used in passive immunisation against bovine
CC mastitis. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 345 AA;

Query Match 63.9%; Score 1727; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 4.1e-87;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 VAADAPAAAGTDTITNQLNTVTVGIDSGTTPVPHQAGYKLYNGFSPVNSAVKGDFTKIVP 258
Db 14 VAADAPAAAGTDTITNQLNTVTVGIDSGTTPVPHQAGYKLYNGFSPVNSAVKGDFTKIVP 73
QY 259 KELNLGVTSKAPPTMAGDQVLAVGIDSGNVITFTDYVNTKDDVKATLTPAYID 318
Db 74 KELNLGVTSKAPPTMAGDQVLAVGIDSGNVITFTDYVNTKDDVKATLTPAYID 133
QY 319 PENVKKTGNVTLATIGSTTANKTVLVDYKFKFYNLSIKGTIDQDKTNNTVQRIYV 378
Db 134 PENVKKTGNVTLATIGSTTANKTVLVDYKFKFYNLSIKGTIDQDKTNNTVQRIYV 193
QY 379 NPSGDNVIAPIVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 438
Db 194 NPSGDNVIAPIVLTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSESYFVNPENFEDVTN 253
QY 439 SVNITFPNPKYKFEPTDQITTPYLVVNGHIDPNSKGLALRSTLYGYNLIWRS 498
Db 254 SVNITFPNPKYKFEPTDQITTPYLVVNGHIDPNSKGLALRSTLYGYNLIWRS 313
QY 499 MSWDNEVAFNNGSGGIDKPKVVPQFDE 528
Db 314 MSWDNEVAFNNGSGGIDKPKVVPQFDE 343

RESULT 15
ABU15854
ID ABU15854 standard; protein; 961 AA.
XX
AC ABU15854;
XX
XX 19-JUN-2003 (first entry)
XX Protein encoded by prokaryotic essential gene #1381.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX *Staphylococcus aureus*.
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR

PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX N-PSDB; ACA19724.

New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 43778; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 961 AA;

Query Match 19.2%; Score 520; DB 6; Length 961;
Best Local Similarity 26.8%; Pred. No. 2.9e-20;
Matches 149; Conservative 89; Mismatches 201; Indels 118; Gaps 15;
QY 1 VGLIGFGLLSKKEADASE---NSVTQSDASNESKNSDSSVSAAKPTDDTNVSDTKTS 57
Db 21 LGTMIYVVGQKEKAAASEQNTTVEESGSSATESKA-----SETQTT 63
QY 58 SNTNNGTSSVAQNPACQETTQSSSTNATTEPTVGTGATTTTNOANTPATTQSSNTNAE 117
Db 64 TNNVNTIDETQSVATSTEQPSKSTQVTTTEAPTQAPKVEETMKS----- 110
QY 118 ELVNQTSNETTFNDNTVSSVNSPQNSTNAENYSTTQDTSTE---ATPSNESAPQSTDA 174
Db 111 -----QEDLPSEKQVADKETTGTQVDIAQPSN----- 136
QY 175 SNKDVVNQAVNTSAPRMR-----AFSLAAVADAPAAAGTDTITNQLNTVTVGID----SG 224
Db 137 -----VSEIKPRMKRSADVTAVSEKVEAEKATGTDVTKVETSSLEGHKNDS 187

```

QY 225 TTVPHQAGYVKLNYGFSVPNSAVKGDTPKTIIVPKELNLNGVTSTAKVPPIMAG--DQVL 282
Db 188 NIIVNPHNAQRVTLKYKWKFGEGIKAGDYDFTLSDNVETHGISTLRKVPBIKSTEDKVM 247
QY 283 ANGVDSGDNVLYTFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKT 342
Db 248 ANGOVINERTIRYTFDYINNKKDLTAEINLNLFIPTTWTQGSQKQVEVTLGQNKVSKE 307
QY 343 VLVDY-----EKYGYFNLSIKGTIDIDIKTNTYRQTIYVNPBGDNVIAPIVLTGNLKEN 397
Db 308 FDIKYLDDGVKDRMG---VTVNGRIDTLNKEEGKFSHFAYVKNQSLTSVTVTGQVTSG 363
QY 398 TDSNALIDQNTSIKVKYVDNAADLSSEYFV---NPENPEDVTNSVNIITFPNPQYKVEF 454
Db 364 YKQSA---NNPTVKYKHIGSDELAESVYAKLDDTSKEEDVTEKVNLSYTSNGGYTLNL 419
QY 455 NTPDQIITTPYIVVNVNGHIDPNKSGDLALRSTLYGYNISNIWR-----SMSWDNEVA 506
Db 420 GDLDN--SKDYVIKYEGEYDQNAK-DLNPRTLSGYHKYPPYPPYPPYPPYPPYPPYPPYPPY 476
QY 507 FNNGSGSGDGIDKPVVP 523
Db 477 FYSNNAKGDKDKPNDP 493

```

Search completed: November 3, 2004, 19:03:02
Job time : 103.005 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 25.894 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VGTLLGFLSSREADASEN.....NGSGSGDIDKPVVPEQDE 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pbp:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pbp:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pbp:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pbp:*
5: /cgn2_6/ptodata/1/iaa/PCFUS COMB.pbp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pbp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2702	100.0	933	3	US-08-293-728-2
2	2702	100.0	933	3	US-09-421-868-2
3	2690	99.6	936	4	US-08-956-171E-5249
4	2690	99.6	936	4	US-08-781-986A-5249
5	1727	63.9	345	3	US-08-856-253-7
6	494.5	18.3	1166	4	US-09-200-650E-7
7	472	17.5	1092	4	US-09-147-405B-15
8	471.5	17.5	930	3	US-09-134-001C-5314
9	471.5	17.5	930	4	US-09-386-962C-10
10	464	17.2	918	4	US-09-200-650E-1
11	452.5	16.7	1027	4	US-08-956-171E-5254
12	452.5	16.7	1027	4	US-08-781-986A-5254
13	441.5	16.3	582	4	US-09-147-405B-13
14	441.5	16.3	593	4	US-09-147-405B-11
15	398	14.7	1315	4	US-09-200-650E-5
16	338	12.5	1155	4	US-09-710-279-1780
17	336	12.4	1742	4	US-09-386-962C-4
18	336	12.4	1742	4	US-09-386-959-4
19	327	12.1	930	4	US-09-200-650E-3
20	285.5	10.6	251	4	US-08-956-171E-5252
21	285.5	10.6	251	4	US-08-781-986A-5252
22	260.5	9.6	556	4	US-09-248-796A-22338
23	252.5	9.3	270	4	US-09-134-000C-3553
24	233.5	8.6	827	4	US-09-248-796A-17307
25	224.5	8.3	2870	4	US-09-479-467A-15
26	224.5	8.3	3178	4	US-09-479-467A-4
27	224	8.3	1112	2	US-08-714-402-2

28	224	8.3	2137	3	US-09-134-001C-4463	Sequence 4463, Ap
29	223	8.3	1161	3	US-09-327-536-2	Sequence 2, Appli
30	210.5	7.8	613	4	US-09-248-796A-17277	Sequence 17277, A
31	207	7.7	712	4	US-09-248-796A-14274	Sequence 14274, A
32	206.5	7.6	2504	4	US-09-328-352-5821	Sequence 5821, Ap
33	205	7.6	1140	4	US-09-538-092-647	Sequence 647, App
34	204	7.5	669	4	US-09-107-532A-6532	Sequence 6532, Ap
35	203	7.5	1721	3	US-08-700-651-5	Sequence 5, Appli
36	203	7.5	1721	3	US-08-928-361B-6	Sequence 6, Appli
37	203	7.5	1721	4	US-09-588-995A-6	Sequence 6, Appli
38	198.5	7.3	542	4	US-09-538-092-289	Sequence 289, App
39	197	7.3	529	4	US-09-248-796A-16703	Sequence 16703, A
40	195.5	7.2	478	4	US-09-134-000C-3709	Sequence 3709, Ap
41	194	7.2	338	4	US-09-538-092-144	Sequence 144, App
42	191.5	7.1	2315	4	US-09-543-681A-5434	Sequence 5434, Ap
43	191	7.1	1837	3	US-08-928-361B-5	Sequence 5, Appli
44	191	7.1	1837	4	US-09-588-995A-5	Sequence 5, Appli
45	190.5	7.1	983	4	US-09-538-092-1320	Sequence 1320, Ap

ALIGNMENTS

RESULT 1
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match	100.0%	Score 2702;	DB 3;	Length 933;
Best Local Similarity	100.0%;	Pred. No. 9.2e-173;		
Matches	528;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
Qy	1	VGTLLGFLSSREADASENSVTQSDASNESKNSDSSVSAAAPKTDITNVSDTKTSNT	60	
Db	23	VGTLLGFLSSREADASENSVTQSDASNESKNSDSSVSAAAPKTDITNVSDTKTSNT	82	
Qy	61	NGETSAQNPAQOETTTQSSSTNATTTETPTVGEATTTTNOANTPATTTQSSNTNABELV	120	
Db	83	NGETSAQNPAQOETTTQSSSTNATTTETPTVGEATTTTNOANTPATTTQSSNTNABELV	142	
Qy	121	NOTSNETTNDNTVSSVNSPQNSSTNAENSVTTQDTSTEATPPSNESAPQSTDSASNDVV	180	
Db	143	NOTSNETTNDNTVSSVNSPQNSSTNAENSVTTQDTSTEATPPSNESAPQSTDSASNDVV	202	
Qy	181	NOAVNTSAPMRAPSLAAVAADAPAGDTITNQLTNTVTGIDSGTTTYPHOAGVVKLYNG	240	
Db	203	NOAVNTSAPMRAPSLAAVAADAPAGDTITNQLTNTVTGIDSGTTTYPHOAGVVKLYNG	262	
Qy	241	FSVNSAVKGDGTFKITVPKELNNGVTSTAKVPPIMAGDQVLANGVDSGNVLYTFTDY	300	
Db	263	FSVNSAVKGDGTFKITVPKELNNGVTSTAKVPPIMAGDQVLANGVDSGNVLYTFTDY	322	
Qy	301	VNTKDDVKALITWPAIDPENVKKTGNVTLATIGTSTANKTVLDVYEKFKFNLSIKG	360	
Db	323	VNTKDDVKALITWPAIDPENVKKTGNVTLATIGTSTANKTVLDVYEKFKFNLSIKG	382	
Qy	361	TIDQIDKNTNTYRTIYVNPFGDNDVIAVLTGNLKPNTDSNALIDQNTSIVKYVDNAA	420	
Db	383	TIDQIDKNTNTYRTIYVNPFGDNDVIAVLTGNLKPNTDSNALIDQNTSIVKYVDNAA	442	

QY 421 DLSEYFVNPFEDVTSNLTFFPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 480
Db 443 DLSEYFVNPFEDVTSNLTFFPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 502
QY 481 LALRSTLYGNSNIIWRSMWDEAFNNGSGDGDIDKPVVPEQDPE 528
Db 503 LALRSTLYGNSNIIWRSMWDEAFNNGSGDGDIDKPVVPEQDPE 550

RESULT 2

US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 100.0%; Score 2702; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 9,2e-173;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 60
Db 23 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOQTQSSSTNATEETPTVGEATTTTNOANTPATQSSNTAEELV 120
Db 83 NNGETSAQNPAQOQTQSSSTNATEETPTVGEATTTTNOANTPATQSSNTAEELV 142
QY 121 NOTSNETTNDNTVSSVNSPONSNAENVTQDTSTTEATPSNNEAPQSDAASKDV 180
Db 143 NOTSNETTNDNTVSSVNSPONSNAENVTQDTSTTEATPSNNEAPQSDAASKDV 202
QY 181 NOAVNTSAPRMRFAFSLAAVAADAPAGTDITNQLNTVTVGIDSGTTVYPHQAGYVKLYNG 240
Db 203 NOAVNTSAPRMRFAFSLAAVAADAPAGTDITNQLNTVTVGIDSGTTVYPHQAGYVKLYNG 262
QY 241 FSVPNKAVKGTFFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTTDY 300
Db 263 FSVPNKAVKGTFFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTTDY 322
QY 301 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKVLDYDYKGFYNLSIKG 360
Db 323 VNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKVLDYDYKGFYNLSIKG 382
QY 361 TTDQIDKNTNTRQTIYVNSPNSGDNVIAPIVLTGNLKNPNTDSNALIDQONTSIKVKYVDNAA 420
Db 383 TTDQIDKNTNTRQTIYVNSPNSGDNVIAPIVLTGNLKNPNTDSNALIDQONTSIKVKYVDNAA 442
QY 421 DLSEYFVNPFEDVTSNLTFFPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 480
Db 443 DLSEYFVNPFEDVTSNLTFFPNQYKVEENTPDDQITTPYIVVNGHIDNSKGD 502
QY 481 LALRSTLYGNSNIIWRSMWDEAFNNGSGDGDIDKPVVPEQDPE 528
Db 503 LALRSTLYGNSNIIWRSMWDEAFNNGSGDGDIDKPVVPEQDPE 550

RESULT 3

US-08-956-171E-5249
; Sequence 5249, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: 882-4871
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5249:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5249:
US-08-956-171E-5249

Query Match 99.6%; Score 2690; DB 4; Length 936;
Best Local Similarity 99.6%; Pred. No. 5.9e-172;
Matches 526; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 60
Db 32 VGTLLGFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKTDITNVSDTKTSNT 91
QY 61 NNGETSAQNPAQOQTQSSSTNATEETPTVGEATTTTNOANTPATQSSNTAEELV 120
Db 92 NNGETSAQNPAQOQTQSSSTNATEETPTVGEATTTTNOANTPATQSSNTAEELV 151
QY 121 NOTSNETTNDNTVSSVNSPONSNAENVTQDTSTTEATPSNNEAPQSDAASKDV 180
Db 152 NOTSNETTNDNTVSSVNSPONSNAENVTQDTSTTEATPSNNEAPQSDAASKDV 211
QY 181 NOAVNTSAPRMRFAFSLAAVAADAPAGTDITNQLNTVTVGIDSGTTVYPHQAGYVKLYNG 240
Db 212 NOAVNTSAPRMRFAFSLAAVAADAPAGTDITNQLNTVTVGIDSGTTVYPHQAGYVKLYNG 271
QY 241 FSVPNKAVKGTFFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTTDY 300
Db 272 FSVPNKAVKGTFFKITVPEKLNGLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTTDY 331

QY 301 VNTKDDVKATLTWPAIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKG 360
DB 332 VNTKDDVKATLTWPAIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKG 391
QY 361 TIDQIDKTNNTYRQTIYVNFSGDNVIAVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAA 420
DB 392 TIDQIDKTNNTYRQTIYVNFSGDNVIAVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAA 451
QY 421 DLSESYFVNPENFEDVTNSVNIIFPNPNQYKVEFTDDQITTPYIVVNGHIDPNKSGD 480
DB 452 DLSESYFVNPENFEDVTNSVNIIFPNPNQYKVEFTDDQITTPYIVVNGHIDPNKSGD 511
QY 481 LALRSTLYGYNSNLIWRSMSWDNEVAFNNGSGDGDIDKPVVPEQDPE 528
DB 512 LALRSTLYGYNSNLIWRSMSWDNEVAFNNGSGDGDIDKPVVPEQDPE 559

RESULT 4
US-08-781-986A-5249
; Sequence 5249, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5249:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5249

Query Match 99.6%; Score 2690; DB 4; Length 936;
Best Local Similarity 99.6%; Pred. No. 5.9e-172;
Matches 526; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VGTLLIGFLLSKKADASENSVTQSDASNESKSNDSVSAAPKTDNTVSDTKTSNT 60
DB 32 VGTLLIGFLLSKKADASENSVTQSDASNESKSNDSVSAAPKTDNTVSDTKTSNT 91
QY 61 NNGETSAQNPAQOBTQSSSTNATTEPTVGEATTTTQANTPATQSSNTNAAELV 120
DB 92 NNGETSAQNPAQOBTQSSSTNATTEPTVGEATTTTQANTPATQSSNTNAAELV 151

QY 121 NOTSNETTFTDNTVSSVNSPQNSTNAENSVTTQDTSSTEATPSNNEAPQSTDSASNDVV 180
DB 152 NOTSNETTFTDNTVSSVNSPQNSTNAENSVTTQDTSSTEATPSNNEAPQSTDSASNDVV 211
QY 181 NQAVNTSAPMRAFSLAAVAADAPAGTDTINQLTNTVTGIDSGTTTYPHQAGVVKLNIG 240
DB 212 NQAVNTSAPMRAFSLAAVAADAPAGTDTINQLTNTVTGIDSGTTTYPHQAGVVKLNIG 271
QY 241 FSPVNSAVKGDTPKTIYVPEKLNNGVTSTAKVPIIMAGDOVLANGVIDSGNVIYTFDY 300
DB 272 FSPVNSAVKGDTPKTIYVPEKLNNGVTSTAKVPIIMAGDOVLANGVIDSGNVIYTFDY 331
QY 301 VNTKDDVKATLTWPAIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKG 360
DB 332 VNTKDDVKATLTWPAIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKG 391
QY 361 TIDQIDKTNNTYRQTIYVNFSGDNVIAVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAA 420
DB 392 TIDQIDKTNNTYRQTIYVNFSGDNVIAVLTGNLKPNTDSNALIDQOQNTSIKVKVDNAA 451
QY 421 DLSESYFVNPENFEDVTNSVNIIFPNPNQYKVEFTDDQITTPYIVVNGHIDPNKSGD 480
DB 452 DLSESYFVNPENFEDVTNSVNIIFPNPNQYKVEFTDDQITTPYIVVNGHIDPNKSGD 511
QY 481 LALRSTLYGYNSNLIWRSMSWDNEVAFNNGSGDGDIDKPVVPEQDPE 528
DB 512 LALRSTLYGYNSNLIWRSMSWDNEVAFNNGSGDGDIDKPVVPEQDPE 559

RESULT 5
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent-in-release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS:

TOPOLOGY: linear
US-08-856-253-7

Query Match
Best Local Similarity 63.9%; Score 1727; DB 3; Length 345;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 VAADAPAAAGTDTINQLTNVTGIDSGTTVYPHQAQVYKLVNPGFSPVNSAVKGDTEKITYP 258
DB 14 VAADAPAAAGTDTINQLTNVTGIDSGTTVYPHQAQVYKLVNPGFSPVNSAVKGDTEKITYP 73
QY 259 KEINLNGVTSKAVPPIMAGDQVLANGVIDSDGNVIYTTFTDYVNTKDDVKATLTPAYID 318
DB 74 KEINLNGVTSKAVPPIMAGDQVLANGVIDSDGNVIYTTFTDYVNTKDDVKATLTPAYID 133
QY 319 PENVKKTGNVTLATGSGTANKTAVLDYKYGKFNLSIKGTIDQIDKTNNTYRTIYV 378
DB 134 PENVKKTGNVTLATGSGTANKTAVLDYKYGKFNLSIKGTIDQIDKTNNTYRTIYV 193
QY 379 NPSGONVIAPVLTKNLPNTDSNALIDQNTSIKVKVDNAADLSSEYFVNPENFEDVTN 438
DB 194 NPSGONVIAPVLTKNLPNTDSNALIDQNTSIKVKVDNAADLSSEYFVNPENFEDVTN 253
QY 439 SVNITFPNPNQYKVEFNTDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNNSIIWRS 498
DB 254 SVNITFPNPNQYKVEFNTDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNNSIIWRS 313
QY 499 MSWDNEVAFNNGSGGDKPKVVPQDPE 528
DB 314 MSWDNEVAFNNGSGGDKPKVVPQDPE 343

RESULT 6
US-09-200-650E-7
; Sequence 7, Application US/09200650E
; Patent No. 6680195
; GENERAL INFORMATION:
; APPLICANT: Patti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A. O.
; APPLICANT: Eidiann, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-7

Query Match
Best Local Similarity 18.3%; Score 494.5; DB 4; Length 1166;
Matches 172; Conservative 91; Mismatches 215; Indels 131; Gaps 27;

QY 1 VGTLLFGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKTDNTNVDTKTSSNT 60
DB 37 VGTLLFG-LGQEAKAENTSTE-NAKQDDATTSNKEV-----VSETENNST 84
QY 61 NNGETSAVNAQAQQTQSSNTNATTEFTVTCGEATTTTNTQANTPATQSSNTNAB-EL 119
DB 85 ENNST----NPIKETNTDSQPEAKKEST-----SSSTOKQNNVTATTETKPKQIEKN 135
QY 120 VNQTSNETTFNDNTV-----SSVNSPQNSTNAENV 150
DB 136 VKPSTDKTATEDTSVILEKKAAPNTNNDVTKPTSEPTSEIQTPTPQESTNIENS 195

QY 151 S-----TTQDTS TEATPSNNE SAP-----QSTDASNKD VVNQAVN 185
DB 196 QROPSPKVDNQVTDATNPKEPVNSKEBKKNPEKELVRNDSNTDHTKPVATPIS 255
QY 186 TSAPRMRAFSLAAVAADAPAAAGTDTINQLTNVT-----VGIDSGTTVYPHQAQVYKLVNPG 240
DB 256 VAPKRVNAKMRFAVAQPAVAASNNV-NDLIKVTQTIKVGDKGNVAAAHADGKDIEYDTE 314
QY 241 FSPVNSAVKGDTEKIL-----TVPKELNNGVTSKAVPPIMAGD-----QVLANGVID-SDG 291
DB 315 FTIDNKKVKGDTMTINYNKXNVPISDL-----TDKNDDPIDITDPSGEVIAKGTDFKATK 367
QY 292 NVIYFTDYVNTKDDVKATLTPAYIDPENV--KKTGNVTLATGISTTANKTAVLDYDEK 349
DB 368 QITYFTDYVNTKDDVKATLTPAYIDPENV--KKTGNVTLATGISTTANKTAVLDYDEK 425
QY 350 YGKFVNLSTIKGTIDQIDKTNNTYRTIYVNP--SGDNVIAVPL-----TGNLKNPTDS 400
DB 426 PMVHGDSNIQSITFTKLDEKQIEQIYVNPPLKKSATNTKVDIAGSOVDYGNILKNGS 485
QY 401 NALIDQNTSIKVKVDNAADLSSEYFV-NPENFEDVTNSVNTFPNPNQYKVEFNTPD- 458
DB 486 -TIID-QNTEIKVYKNSDQQLPQSNRIYDFSOYEDVTSQ-----FUNKKSFNNVAILDF 539
QY 459 DOITTPYIVVNGHIDPNKSGDL-----ALRST-LYGYNNSIIWRSMSWDNEVAFNNGS 511
DB 540 GDINSAYIIKVKVSKYPTSDGELDTAAGTSMETTDKYGK-----YNYAGYSNFTVTSNDT 594
QY 512 GSGDGDIDKP 520
DB 595 GGGDGTVPK 603

RESULT 7
US-09-147-405B-15
; Sequence 15, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingemar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09/147,405B
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1092
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-15

Query Match
Best Local Similarity 17.5%; Score 472; DB 4; Length 1092;
Matches 157; Conservative 103; Mismatches 238; Indels 94; Gaps 25;

QY 1 VGTLLFGLLSKEADASENSV-----TQSDASNESKNSDSSVSAAPKTD 48
DB 36 IGATLLFG-LGNEAKAEENSVDYKDSNTDDELSDNSDQSDDEKNDVNNQNSINTDD 94
QY 49 TWV-----SDTKTSNTN--NGTSTVAQNP-----AQQTTSSTN 83
DB 95 NNQIITKRETNNDYGLIEKRESDRTSTTNVDNEATFLOKTDNTHLTETEEVKSSSVE 154


```
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-13

Query Match          16.3%; Score 441.5; DB 4; Length 582;
Best Local Similarity 26.3%; Pred. No. 1e-21;
Matches 145; Conservative 99; Mismatches 227; Indels 81; Gaps 23;

QY 29 SNESKNDSSSSVAAPKTDITNV-----SDTKTSNTN---NGETSAQ 69
Db 1 SDEEKNDVINNNQSDINTDNNQIIKKEETNNYDGEKRSDETRSTTNVDENEATFLOK 60

QY 70 NP-----AQQTSSSTNATTEETPVTEGATTTTNOANTPATQOS--SNTNAELVN- 121
Db 61 TPQDNTHLTREEVKESSSVSSSSIDTAQQSHHTINREESVQSDNVEDSHVDFANS 120

QY 122 ---QTSNETTFNDITNV-----SSVNSPQNSTNAENVSTQDTSTPATPSN---NES 167
Db 121 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELLN-LPINEYENKA 179

QY 168 APOQSDASNKDVVNOAVNTISAPRMAFSLAAVAADAPAAAGTDTITNOLTNVTVGI--DSGTT 226
Db 180 RPLSTTSQAQPSIKRVTNV-----QLAAEQGSNNVHLIKVTDQ--SITEGYDDSEGV 228

QY 227 VYPHQAQVYKLVNFGYSPVNSAVKGTFTKITVPKELNGLNGVTSTAKVPPIMAGD-OVLANG 285
Db 229 IKAHDAENLIYDVTVEVDKVKSGDTMTVDIDKNTVPSDLTDSFIPKIKDNSGELIATG 288

QY 286 VIDS--DGNVIYTFDYVNTKDDVKATLMPAVIDPENVKKTG---NVLATGIGSTANK 341
Db 289 TYDNKNKQITFTFDYVDKYENIKAHKLKITSYIDSKVPNNNTKLDVEYKTALES--VNK 346

QY 342 TVLVDEYKYGKYNLSIKGTIDQIDKNTNTYQTTIYVNPBGDNVIAPLVLTGNLKNPTDSEN 401
Db 347 TITVBYQPNENRNTANLOSMTFNIDTKNHTVEQTIYINPL--RYSAKETNNVINSNGDEG 404

QY 402 ALIDQOQNTSIKVKYVDNAADLSESYFV-NPENFEDVTNSVNITFPNPGYKVEFNTPDQ 460
Db 405 STIIDDSIIKVKYVDNQNLPSNRIDYSEYEDVTNDYQAQLGNNNDVNI--N 460

QY 461 ITTPYIVVNGHIDPNSKGD-----LALRSTLYGYSNIIWRSMWMDNEVAFNNGSGS 513
Db 461 IDSPYIIKVISKYDPN-KDDYTTIQOQVTMTQTINEYTGEE--FRTASYDNTIAFSTSSGQ 517

QY 514 GDGIDKPVVPEQ 525
Db 518 SQG-DLP--PEK 526

RESULT 14
US-09-147-405B-11
; Sequence 11, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: Coagulase-Negative Staphylococcus
; CURRENT APPLICATION NUMBER: US/09147405
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
```

```
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-11

Query Match          16.3%; Score 441.5; DB 4; Length 593;
Best Local Similarity 26.3%; Pred. No. 1.1e-21;
Matches 145; Conservative 99; Mismatches 227; Indels 81; Gaps 23;

QY 29 SNESKNDSSSSVAAPKTDITNV-----SDTKTSNTN---NGETSAQ 69
Db 8 SDEEKNDVINNNQSDINTDNNQIIKKEETNNYDGEKRSDETRSTTNVDENEATFLOK 67

QY 70 NP-----AQQTSSSTNATTEETPVTEGATTTTNOANTPATQOS--SNTNAELVN- 121
Db 68 TPQDNTHLTREEVKESSSVSSSSIDTAQQSHHTINREESVQSDNVEDSHVDFANS 127

QY 122 ---QTSNETTFNDITNV-----SSVNSPQNSTNAENVSTQDTSTPATPSN---NES 167
Db 128 KIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDELLN-LPINEYENKA 186

QY 168 APOQSDASNKDVVNOAVNTISAPRMAFSLAAVAADAPAAAGTDTITNOLTNVTVGI--DSGTT 226
Db 187 RPLSTTSQAQPSIKRVTNV-----QLAAEQGSNNVHLIKVTDQ--SITEGYDDSEGV 235

QY 227 VYPHQAQVYKLVNFGYSPVNSAVKGTFTKITVPKELNGLNGVTSTAKVPPIMAGD-OVLANG 285
Db 236 IKAHDAENLIYDVTVEVDKVKSGDTMTVDIDKNTVPSDLTDSFIPKIKDNSGELIATG 295

QY 286 VIDS--DGNVIYTFDYVNTKDDVKATLMPAVIDPENVKKTG---NVLATGIGSTANK 341
Db 296 TYDNKNKQITFTFDYVDKYENIKAHKLKITSYIDSKVPNNNTKLDVEYKTALES--VNK 353

QY 342 TVLVDEYKYGKYNLSIKGTIDQIDKNTNTYQTTIYVNPBGDNVIAPLVLTGNLKNPTDSEN 401
Db 354 TITVBYQPNENRNTANLOSMTFNIDTKNHTVEQTIYINPL--RYSAKETNNVINSNGDEG 411

QY 402 ALIDQOQNTSIKVKYVDNAADLSESYFV-NPENFEDVTNSVNITFPNPGYKVEFNTPDQ 460
Db 412 STIIDDSIIKVKYVDNQNLPSNRIDYSEYEDVTNDYQAQLGNNNDVNI--N 467

QY 461 ITTPYIVVNGHIDPNSKGD-----LALRSTLYGYSNIIWRSMWMDNEVAFNNGSGS 513
Db 468 IDSPYIIKVISKYDPN-KDDYTTIQOQVTMTQTINEYTGEE--FRTASYDNTIAFSTSSGQ 524

QY 514 GDGIDKPVVPEQ 525
Db 525 SQG-DLP--PEK 533

RESULT 15
US-09-200-650E-5
; Sequence 5, Application US/09200650E
; Patent No. 6680195
; GENERAL INFORMATION:
; APPLICANT: Fatti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eidehinn, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
```


; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-5

Query Match 14.7%; Score 398; DB 4; Length 1315;
Best Local Similarity 24.5%; Pred. No. 2.5e-18;
Matches 135; Conservative 108; Mismatches 252; Indels 56; Gaps 19;

QY	1	VGTILIFGL--LSSKEADASENSVTQDSASNESKSDSSVSAAPKTDIT--NVSDTKTS	57
DB	37	VGTILIFGLNQEAKEAESTINKELNEATTSASDNQSSDKVDMQQLNQEDNTKNDNQKEMV	96
QY	58	SNTNGETSVQAQPAQOETQSSSTNAATTEETPTVGEATTTTNTQANTPATQSSNTNAE	117
DB	97	SSQGNETTSGNKLIEKESVQSTTGKV-----EVSTAKSDEQASPKSTNEDLNTKQ	148
QY	118	ELVNOTSNETFTFNTVTNVSSV--NSPQNSINAENVSTTQDTSTEATPSNESA-----PQ	170
DB	149	TISNQEALQPLQENKGVVNVQPTNEENKKVDKTESITLNVKSDAIXSNDETLVDNNSN	208
QY	171	STDASNKDVVNOAVNTSAP-----RMRAFSLAAVAADAPAAAGTDITNQLTNVTVIGDSGT	225
DB	209	SNNENNADII--LPKSTAPKELNTRMRIAAVQPSSTEAKNVNDLITSTTILTVVDADKNN	266
QY	226	TVYPHQAGYKLVNGFSPNSAVKGDFTKIVPKELNLNGVT-----STAKVPPIMAG	278
DB	267	KIVPAQ--DYLSLKSGQITVDDKVKSGDYFTIKYSDTVQVYGLNPEDIKNIGDIKDP---NN	322
QY	279	DQVLANGVIDSDGVI-YTFTDYVNTKDDVKATLTMPAYIDPENVKKTGN-VTLATGIGS	336
DB	323	GETIATAKHDTANNLIITYFTDYVDRNSVQMGINSIYMDADTIPVSKNDVEFNVTIGN	382
QY	337	TTANKTVLVDYKYGKFNLSIKG---TIDQIDKTNNT--YRQTIYVNPSPGDNVIAPVL	390
DB	383	TTTTKTANIQYPDYVNVNEKNSIGSAFTETVSHVGNKENPGYKQTIYVNPSENSLTNAKL	442
QY	391	TGNLKPNTDSNAL--IDQNTSIKVKYVDNAADLSESYFVNPNFEDVTSV--NITFPN	446
DB	443	KVQAYHSSYPNNIGQINKDVTDIKIYQVPKGYTLNKGVDVNTKELTDVTNOYLQKITYGD	502
QY	447	PNQYKVENTPDDQITTPYIVVNGHID-PNSKGDALRLSTLYGYSNIIWRSMWDNEV	505
DB	503	NNSAVIDFGNAD----SAYVVMVNTKFQYTNSESPTLVQMATLSSTGN---KSVSTGNAL	555
QY	506	AFNNGSGSGDG	516
DB	556	GFTNNQSGAG	566

Search completed: November 3, 2004, 19:11:51
Job time : 27.894 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 19:03:11 ; Search time 7.73844 Seconds
(without alignments)
1078.240 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VOTLIGFLLSKREADSEN.....NGSGSDGIDKFPVPEQDPDE 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 66696 seqs, 15802848 residues

Total number of hits satisfying chosen parameters: 66696

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	7.4	1087	6	US-10-967-702-299
2	198	7.3	503	8	US-60-613-194-170
3	192.5	7.1	2165	6	US-10-732-923-13547
4	183	6.8	1584	6	US-10-732-923-13548
5	179.5	6.6	1430	6	US-10-797-821-36
6	168	6.2	3328	6	US-10-732-923-8311
7	165	6.1	1903	1	PCT-US04-02460-3
8	163.5	6.1	2159	6	US-10-732-923-9917
9	162.5	6.0	4315	6	US-10-399-103A-811
10	159.5	5.9	1849	6	US-10-819-275-2
11	159	5.9	3147	6	US-10-732-923-10253
12	159	5.9	3167	6	US-10-732-923-10252
13	157.5	5.8	2964	6	US-10-732-923-13552
14	156.5	5.8	585	6	US-10-868-381-58
15	155	5.7	449	6	US-10-955-952-224
16	155	5.7	449	6	US-10-157-779-224
17	155	5.7	449	6	US-10-964-241-224
18	153.5	5.7	637	6	US-10-732-923-928
19	153.5	5.7	1457	6	US-10-915-740A-1047
20	152.5	5.6	793	6	US-10-732-923-3091
21	152	5.6	1554	6	US-10-797-821-38
22	151.5	5.6	2719	6	US-10-732-923-8668
23	151	5.6	1475	6	US-10-797-821-34
24	150.5	5.6	3664	6	US-10-967-702-79
25	150	5.6	1375	6	US-10-797-821-35

26	150	5.6	1404	6	US-10-732-923-4303	Sequence 4303, Ap
27	147.5	5.5	1063	6	US-10-732-923-8595	Sequence 8595, Ap
28	146.5	5.4	585	6	US-10-868-381-57	Sequence 57, Appl
29	145.5	5.4	585	6	US-10-868-381-56	Sequence 56, Appl
30	145	5.4	1627	6	US-10-819-275-6	Sequence 6, Appl
31	145	5.4	2541	6	US-10-732-923-8761	Sequence 8761, Ap
32	144.5	5.3	686	6	US-10-732-923-935	Sequence 935, App
33	144.5	5.3	1026	6	US-10-478-676A-1	Sequence 1, Appl
34	143.5	5.3	1985	6	US-10-732-923-3351	Sequence 3351, Ap
35	143.5	5.3	6761	6	US-10-732-923-15035	Sequence 15035, A
36	142.5	5.3	1576	6	US-10-732-923-19033	Sequence 19033, A
37	142.5	5.3	2271	6	US-10-732-923-19032	Sequence 19032, A
38	142.5	5.3	2307	6	US-10-732-923-19147	Sequence 19147, A
39	142.5	5.3	2314	6	US-10-732-923-19146	Sequence 19146, A
40	142.5	5.3	4226	6	US-10-732-923-22586	Sequence 22586, A
41	142.5	5.3	4226	6	US-10-732-923-22707	Sequence 22707, A
42	142	5.3	530	8	US-60-613-207-153	Sequence 153, App
43	142	5.3	533	8	US-60-613-207-159	Sequence 159, App
44	141	5.2	419	8	US-60-613-207-151	Sequence 151, App
45	141	5.2	422	8	US-60-613-207-147	Sequence 147, App

ALIGNMENTS

RESULT 1
US-10-967-702-299
; Sequence 299, Application US/10967702
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Pauloski, Nicole
; APPLICANT: Taylor, Ian
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: GENE EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 5176
; CURRENT APPLICATION NUMBER: US/10/967,702
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/508,355
; PRIOR FILING DATE: 2003-10-03
; NUMBER OF SEQ ID NOS: 400
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 299
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-967-702-299

Query Match	7.4%	Score 199;	DB 6;	Length 1087;
Best Local Similarity	21.7%	Pred. No. 4.1e-05;		
Matches 127;	Conservative 99;	Mismatches 236;	Indels 124;	Gaps 27;
Qy	11	SSKEADASENSVTQSDASN-ESKNDSSSVSAAPKTDITNVSDT-----KTSNTN	61	
Db	541	SSSQAPSSLLTSTASESSSTISSNQESGYQSGPIQSTTYTSONNAQGLYQQRSTQR	600	
Qy	62	NGTSTVAQNPQAQBTQOS-----SSTATTETPTVGEATTTTNTQANTPATTOSSNTNAE	117	
Db	601	RYPSSTSSSP-QKDLQAKNGFSSVQATQQTQSVGATGSAVKSDSPST--SSIFPLN	657	
Qy	118	ELVNQTSNETTFNDNTVSVNSPQNSNAENVSTTQDTSTTEATPNN-ESAPOSTASN	176	
Db	658	ETVSAASALLTTTQHS--SSLGGLSHSEEPNTTTTQHSSTLSQOQNTLSSTSSGRTST	715	
Qy	177	KDVVNQAVNTSA---PMRAFSLAAVAADAPAGTDITNOL---TNVTVGIDSGTTVYPH	230	
Db	716	STLTLTVESEANLHSSSTFTSTSTVSVAPPVSVSSSLNSGSLGLSGNSTVTAS	775	
Qy	231	QAGVVKLNYGFSVPNSAVKGDFTKITVPEKLNGLNVGTSTAKVPPIMAGDOVLANGVIDSD	290	
Db	776	TRSVATTSKAPN-----LP-----FCVPLEPNPYIMAPGLLHAY	813	
Qy	291	GNVIYTFDYVNTKDDVKATLT-MPAYIDPENVKKTGNVTLATIGTGTANKTVLVDEYK	349	

Db 814 PFQVGY-----DDLQMLQTRFP--LDYYSIPFPPTPTPLTGRDGLASNPYSGDLTK 864
QY 350 YGK--FYNLISIKGTIDQID---KTNNTYRQIYVNPSS---GDNVIA-PVLIGNLKENPD 399
Db 865 FREGDASSPAPATTLAQPOONQTOHTHTTQOT-FLNPALPFGYSYLSLYTG--VPGLP 921
QY 400 SN-----ALIDQNTSIVKYVDNAADLSYFVNPF--EDVTSNVNITFPN-----446
Db 922 STFYCPAVFPVAPTSSKGHGVNVSNASATPFQOPSGYSGHYNTGVSTSNTPVPI 981
QY 447 -----PNQYKVE-----FNTDDQITTPYIVVNGHIDPNKSGDLALRSTLYG-----490
Db 982 SGVSYSKTSQSEKQFSGTGAAGFNPALGSGGPINP-----ATAAAYPPAPF 1032
QY 491 -----NSNIITWRMSWNEVAFNNGSGDGDIDKPVVPEOP 526
Db 1033 MHILTPHQPHSQILLHHLQDQG-----TSGQBSQTSSIPQKP 1072
RESULT 2
US-60-613-194-170
; Sequence 170, Application US/60613194
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING PROTEINS ASSOCIATED WITH ABIOTIC
; TITLE OF INVENTION: STRESS RESPONSE AND PLANT CELLS AND PLANTS WITH INCREASED
; TITLE OF INVENTION: TOLERANCE TO ENVIRONMENTAL STRESS
; FILE REFERENCE: AE2004-0486
; CURRENT APPLICATION NUMBER: US/60/613,194
; CURRENT FILING DATE: 2004-09-24
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 170
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-60-613-194-170

Query Match 7.3%; Score 198; DB 8; Length 503;
Best Local Similarity 23.1%; Pred. No. 1.9e-05;
Matches 98; Conservative 76; Mismatches 181; Indels 70; Gaps 15;
QY 17 ASENSVTSQDSASNESKNDSSVSAAPKTD-DTNVSDTKTSNTNNGTSVAQNPAQOE 75
Db 120 ASTADSTSTATSTSTSSSTSVSSKTSKLDTKTSSGATHSSSSSTSTSTSSSE 179
QY 76 TQSSSTWATTEETPVTCGATT--TTNQAATPATTQSSNTNABELVNQTSNETFDN 133
Db 180 TTTSSSSSSSSSTSTSTSTSSSTSTSSSTSTSSSTSTSSSTSTSSSTSTSSST 238
QY 134 TVSSVNSPQNSTNAENVSTTQDTSTEATPSNNEAPQSTASNKDVNQAQVNTSAPRM-- 191
Db 239 STSS-----STSTATVTSFSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 286
QY 192 RAPSAAVAADAPAGTDITN-----QLTNVTVGIDSGTTVPHQAGYVKL 237
Db 287 TRTSVYATVTSSTSSSTSSLLKSSSKSGLSGAGIAGVVGVCGT-----VALLAL 340
QY 238 NYGFSV-----ENSAYKGTETKITVPEKLNANGVTSTAKVPPIMAGDQVLANGVDSGNV 293
Db 341 ALFFVWKRRQSSQSHVDLETKQYQYSLGDAANPVVPSASS-----TNWHIPSRNNT 396
QY 294 IYTFDYVNTKDDVKATITMPAYIDPENVKKTGNVTATGIGSTTANKTVL--VDYKYG 351
Db 397 ALS-----KNTASTATVLDLPTRAPGGRDSIITGDAHNISKSHFSPVVEBPP 445
QY 352 KFYN-----LSIKGTIDQIDKTNNTYRQIYVNPSSDNIIVAPVLIGNLKPDNLSALIDQ 407
Db 446 SIYNGORFATSPLDMMEE-----RQLHIVNP--DNVSSNIGSNVSGDDDDYDDAKDSN 498
QY 408 NTSIK 412
Db 499 NSSLR 503

RESULT 3

US-10-732-923-13547
; Sequence 13547, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 13547
; LENGTH: 2165
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-13547

Query Match 7.1%; Score 192.5; DB 6; Length 2165;
Best Local Similarity 18.6%; Pred. No. 0.00024;
Matches 97; Conservative 76; Mismatches 146; Indels 203; Gaps 20;
QY 18 SENSVTQSDASNESKNDSSVSAAPKTDTDNVSDDTKTSSNTNNGTSVAQNPAQOETT 77
Db 1090 SKSSINDSSSSNNNNNN-----TTTINDSASTKGNNN---EISSPETYQDQV 1138
QY 78 QS-----SSTNA-----TTEETPVTEGATTTTNOANTPATQSSNTNABELVNQT 123
Db 1139 KSKINFYBHVNISSITNLNNVNTSTNPTNPSNVNQPSNITTATTATTSTNNNNVNS 1198
QY 124 SNETTFDNTVSSVNSPQNSTNAENVSTTQDTSTEATPSNNEAPQSTASNKDVNQA 183
Db 1199 INNNNNNN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNILSKE 1253
QY 184 VNTSAPRMRAFSLAAVAADAPAGTDITNQ--LTNVTGIDSGTTVYPHQAGYVKLVNGF 241
Db 1254 NSTNS-----LNNLLNNTSVG-----1271
QY 242 SVPSNAVKGDTFKITVPKELNMGVTSTAKVPPIMAGDQVLANGVIDSGNVIYT-----296
Db 1272 KTHNRSSSG-----SDSIQPPPLPTG-----GSSHNIIFYSAFPTH 1306
QY 297 -FTYVNTKDDVKATLTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVDYKYGKFN 355
Db 1307 PYTD-----IDQSTYTRT-----FGSSINNRKSLPPEYTNQVLN 1343
QY 356 LSIKGTIDQIDKTNNTYRQIYVNPSSDNIIVAPV-----LTGNLKPNTDS 400
Db 1344 F-----LAHSNN 1398
QY 401 NALIDQNTSKYKVDNAADLSYFVNPFENFEDVTSNVNITFPNPNQYKVEFNTPDQ 460
Db 1399 NN-----DSFSDINDNNNSVVGNDPEQ-----DDQ 1422
QY 461 I-----TTPYIVVNGHIDPNKSGDLA-----LRSTL 487
Db 1423 ILQNGISTTSTIVPTN-----DETQKLSQLENSKNKVRSTV 1460

RESULT 4

US-10-732-923-13548
; Sequence 13548, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149

```
; SEQ ID NO 13548
; LENGTH: 1584
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-13548

Query Match          6.8%; Score 183; DB 6; Length 1584;
Best Local Similarity 18.4%; Pred. No. 0.00062;
Matches 121; Conservative 107; Mismatches 229; Indels 202; Gaps 31;

QY 16 DASENVTSQDSASNEKSDSSVSAAPKTDN--VSDTKTSNNGETSVAQNPQAQ 73
Db 409 NSNNNNNNNNIIGKKTITTTTSTPSSINNEDISSNNNNNNNNNNNNNNNN 468
QY 74 QETQSSNTVAITEETPVTEATTNTTQANTPAITQSSNTNAEELVNQTSNETTENDTN 133
Db 469 NNNNNNNNNNSNNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 524
QY 134 TVSSVNSPQNSTNAENVSTTQDTSSTATPNNESAPQSTDA---SNKVQVQAV--NT 186
Db 525 NNNNNNNNNIYTKKPSIGTDESSTGLGNNSSGNNSSGSGNNSIIKORSPPHSI 584
QY 187 SAPRMAFSLAAVAADAPAGTDTNQL---TNVTVGIDSGTTPVPHQAGYVKLN---Y 239
Db 585 NGPLM-----LPPSSITNNNNIYSSYNSTTAG--SSTILP-----TLNHPIF 625
QY 240 GFSVPNS-----AVKGD-----TFKITVPKELNLN---GVSTAKVPIMAGDQV 281
Db 626 GNTTNNSSSTLSVGGNNLLGRHCQSLPITASTNHTLSSSLGVSPSPSPKTSRPRK 685
QY 282 LANGVIDSGNVLYTFDY-----VNTKDDVKAILTMPAYIDPENVKKTGN 327
Db 686 IVNS--SEDLGFVQTQDQGPSPAWRRCKGSIKTKDDI--TLTI-----IKKTS 733
QY 328 VTLAGTIGSTANKTV-----LVDEYKYG-----KFYNLSIKG 360
Db 734 VAMADRPSSNSSTTICFEVYLEGHDKKGSIWGLSHSTPYPIKHIGREPKSYGFSSEG 793
QY 361 TIOQIDKNTTYQTYYVNPDSGNVIAPVLTGNLKPNTSDNALIDQON-----T 409
Db 794 EYKGGSEIGEPYGPFFFFD--GDSIASSCVIG--CGINTSTRDIFTKNGHYLGVAFSRVT 850
QY 410 SIKY-----KVDNAADLSESYF---VNPE---NFED-VT 437
Db 851 SDLYPISIPRGVVGGLCVATPFGCHFRFNIEDLPGLSPSWTEALGPDQGGSGFKWAP 910
QY 438 NSVNIITPP--NPQYKVEFTPDQITTPYVVVNGHIDPNSKGDALRLSTLYGYSNI- 494
Db 911 NDVAIMLESFNYGYRKNFR--DNNISGRHL---EGITHAMKNDLGIET--YGHREDII 963
QY 495 -----IWRSSWDN-----EYAFNNGSGSGDGI 517
Db 964 NRLNRMITQIWNDRKSPDYPYKIAIDSSDKIRWPASGGSGGINISGGVYIGSSSGSDGI 1022

RESULT 5
US-10-797-821-36
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-36

Query Match          6.6%; Score 179.5; DB 6; Length 1430;
Best Local Similarity 19.3%; Pred. No. 0.00091;
Matches 106; Conservative 80; Mismatches 211; Indels 153; Gaps 21;

QY 3 TLIGFLLSKRADASENVTSQDSASNEKSKNDSSVSAAPKTDNTVSDTKTSNTNN 62
Db 31 TTLIGSSVSAETEQQTSKQVVTQK--SEDDKAASESSQTDAPKTKQAQTEQTOAQANV 87
QY 63 GETSVA--QNPAAQOQETTSSTNATTEETPVTEATTNTTQANTPAITQSSNTNAEEL 119
Db 88 ADTSITITKETSQNTTQANSDDKTV-----TNTKSEEA 122
QY 120 VNQTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDTSSTATPNNESAPQSTDAKNQV 179
Db 123 --QTSEERT-----KQSEEAQTTASSQALTO-AKAEITKQROTAAQE-----NKNP 165
QY 180 VNQAVNTSAPRMAFSLAAVAADAPAGTDTIINQLTNVTVGIDSGTTPVPHQAGYVKLN 239
Db 166 VDLA---AIPNVKQI-----DGKYYIGSDGQPKKNF 194
QY 240 GFSVPNSAVKGTFFKITVPKELNLNGVTSTAKVP-----PIMAGDQVLANGVIDSDGNVI 294
Db 195 ALTVNNKNVLYFDK-----NTGALTDTSQYQFQGLTKUNDDTYPHQNIVNFENTSL 245
QY 295 YTFDVTNKTDDVKAILTMPAYIDPENVKKTGNVTLAGT-----IGSTTANKTVLVDY 347
Db 246 ETIDNV-----TADSWYRPKDLKNGKTWTASSSEDLRPLLMSSWWPDKQTIAY 295
QY 348 EKY-----GKFYNLSIKGTIDQIDKNTTYQTYYVNPDSGNVIAPVLTGNLKPNTD 399
Db 296 LNYMNOQGLGTGENY-----TADSSQESLNLAQTVOVK-----IETKISQTOQ 339
QY 400 SNALLIQOQNTSIKVKYKVDNAADLSESYFVNPENFED-----VTNSVNITPPNPQYKVEFN 455
Db 340 TQWRDLINSFVQTQNNWNSQTESDTSAGEKHDLGGALLYSNDDKTAYAN--SDYRLNLR 398
QY 456 TPDDQITTPYVVVNGHIDPNSKG--DLALRLSTLYGYSNIIWRSSWDN-----DNEVAF 507
Db 399 TPTSQTKCP-----KYFEDNSSGGYDFLLANDIDNSNPVVQAEQLNWLHLYMNYGSIVA 452
QY 508 NNGSGSGDGI 517
Db 453 NDPEANPDGV 462

RESULT 6
US-10-732-923-8311
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8311
; LENGTH: 3328
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
```


Qy	76	TTOSSSNNAATTEP	-----VTGEATTT-----TTNOANTPAT-----TQS	111
		::: : :	: : : : :	:
Db	2925	TTATISSTSPSSPTGTTWILTELTAAATTAGTGPATPSSPTGTTWILTELTATTT	2984	
Qy	112	SNTNAEBLVQTSNET--TFNDTNTVSSVNGSPQNST-----NAENVSTTQD----	T	156
		: : : : :	: : : : :	:
Db	2985	ASTGSTATLSSTPGTTWILTEPSTATVTPPQSTATASSQATAGTAPHVSTTATTTVT	3044	
Qy	157	STRATPSSNPFASQSTDAASNKDVNVQAVNTSAPMRAPFLAAVAADAPAAAGTDIT--	NQL	214
		: : : : :	: : : : :	:
Db	3045	SSKATPS-----SSPGTATAL-----PALRSTATTPTATSTAIPTSS--SLGTTWRLSQ	T	3093
Qy	215	TNTVTGIDSGT-----TVYPHOAGVVKLVNGFVSYPNSAVKGDTFKLTVPKELNL	263	
		: : : : :	: : : : :	:
Db	3094	TPTATNMSTATPSSPTETVHTSTVLTATTGATGAVATP--SSTPGTAHTTKVP--TTTT	3151	
Qy	264	NGVTS-----TAKVPPPIMAGQVLANGVIDSDGNVIYFTDVTNKKDDVKATLT--	M	313
		: : : : :	: : : : :	:
Db	3152	TGPTATPSSSPGATLTPVW-----I--STTPTPTTPTTSGSVTPSSI	3195	
Qy	314	PAYIDPENVKKTGNVTLAGTGSTANKTVLVDYKYGKFPNLSIKGTIDDKNTNMYR	373	
		: : : : :	: : : : :	:
Db	3196	PGTHTARVLTTTTVATGSMATPSSST-----QTSCTPPPSLATTATTT	3241	
Qy	374	QT-IYVNPS--GDNVIAPVLT	391	
		: : : : :	: : : : :	:
Db	3242	ATGSTNFPSSPTGTTTPPVLT	3263	
RESULT	10			
US-10-819-275-2				
;	Sequence 2, Application	US/10819275		
;	GENERAL INFORMATION:			
;	APPLICANT: Chr. Hansen A/S			
;	TITLE OF INVENTION: Composition with heart rate reducing properties			
;	FILE REFERENCE: P1040FC00			
;	CURRENT APPLICATION NUMBER: US/10/819,275			
;	CURRENT FILING DATE: 2004-04-07			
;	NUMBER OF SEQ ID NOS: 18			
;	SOFTWARE: PatentIn version 3.1			
;	SEQ ID NO 2			
;	LENGTH: 1849			
;	TYPE: PRT			
;	ORGANISM: Lactobacillus helveticus			
US-10-819-275-2				

```

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPOLYNUCLEOTIDES
; FILE REFERENCE: 21272-115/ 785
; CURRENT APPLICATION NUMBER: US/10/399,103A
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: PCT/US01/27760
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 09/687,527
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 892
; SOFTWARE: Custom
; SEQ ID NO 811
; LENGTH: 4315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-399-103A-811

Query Match          6.0%; Score 162.5; DB 6; Length 4315;
Best Local Similarity 22.1%; Pred.No. 0.038;
Matches 111; Conservative 57; Mismatches 163; Indels 171; Gaps 22;

Qy      18 SENVTSQDSASNEKNDSSS-----VSAAPKTDNVDNVDKTSN----- 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2805 TELTTTATTAGTGTATGTPSSSTPGTAPPKVLTPATTTATSSKATSSSPRTATLPV 2864
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      60 -----TNGEISVQAQNAQ-----QE 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2865 LTSATKATSTATSVTPIPSSLTGTTGLPEQTTPVATMTSHPSSTPETHTSTVLTTKA 2924
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 549 KSVIVQALNNSARQTLMSDFTSYG---PTSSL---AFKPDISAPGGHIMSTQNNNGYT 602
QY 268 S---TAKVPPIMAGDOVLANGVI--DSDGNVIYTF-----TDVNTKDDVKATLTWPA 315
Db 603 NMSGTSMASPFATGAQALVQSMNDKNGAFYATYQKMSABERTPPFKITL---MNTAS:QP 660
QY 316 YIDPENV-----KKTGNVTLATGIGSTTANKTVLVDYKGYKFNLSIKG-TID-QIDK 367
Db 661 DISHNVIVPRQAGFINANATIALAKNPSVTVSSNGYPGVELKFKDRLNFQVKF 720
QY 368 TNNYRQTIY--VNPBGDNIAPVLGTNLKPNNTDSNALIDQQNTSIKVKVDNAA--DLSES 425
Db 721 TNRNKALTYKLANNGK--SDVYTS---ATDSSAVLYDK-----KIDGASVKASGD 767
QY 426 YFVNPENEDVTVNSNITFPN---ENQYK---VEENTPD-DQITTPYIVVNGHIDPNK 478
Db 768 IFVPANSTKELT--LTLTLPDFKENQYVEGELTFENSSDSSQLRPLYM-----GFFGDWAS 821
QY 479 GDALRSTL 487
Db 822 SDSLPIFASL 830

RESULT 11

US-10-732-923-10253
; Sequence 10253, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10253
; LENGTH: 3147
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-732-923-10253

Query Match 5.9%; Score 159; DB 6; Length 3147;
Best Local Similarity 20.1%; Pred. No. 0.043;
Matches 123; Conservative 97; Mismatches 203; Indels 190; Gaps 30;

QY 12 SKEADASENSVTQSDSASNESKSNDSVSAAPKTD--TNVSDTKTSSN-TNNGETSVA 68
Db 1931 TKQEKQORSVVVETSPSNKHIS-DESSIS---MDEIFSRQDNKSTNFKSGSIPIL 1985
QY 69 ONPAQOQETQSSSTNATTEPTVGEATTT-----TNQA----- 103
Db 1986 VLPGEKEVASASIN--LNGVLEGGKQKSTDYGREKVKQVENSYSYSTRHMDGASIS 2043
QY 104 -----NTPATQSSNT---NAEE-----LVNQTNETTFDNTVSSVNSPQNSTN 146
Db 2044 LDDIFNTSSTQKTETKIDNSQEPQLSKPVLKSSI LDDLFNNTSGI-----EKT 2096
QY 147 AENVSTTQDTS TEATPSNNEAPQSDASNKDVVQAVN-----TSAPRMRAFLAAVAA 201
Db 2097 SEKTTTTTTTTRTET-----TDYVKRSTSLVDRFGYETATP--PAISIAAISF 2144
QY 202 DAPAAAGTDITNQLTNVTGIDSGTTVYP---HQAGYVKLVNGYFVNSAVKGDTFKITVP 258
Db 2145 DQPS-----SSQAFFRSKHQ-----NLSSLTVP 2169
QY 259 KELNUNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYFTDYVNTKDDVKATLTWPA--Y 316
Db 2170 GKWNESMMSNTSTI-----SLDSDFNNSFSKNTSQV-----EPRMRKPLTLFVDNW 2217
QY 317 IDPENVKKTGNVTATGIGSTTANKTVLVDYKGYKF---YNLSIK-----GTIDQ- 364
Db 2218 ID-----NLVSEATNEATKEAPTKPSDNTLNYFRSPTRISQEIKEYWADMIGDIRK 2270
QY 365 -----IDKTNNTYRQTIYVNPBGDNIAPVLGTNLKPNNTDSNALIDQQNTSIK 413
Db 2271 NMSGTSMASPFATGAQALVQSMNDKNGAFYATYQKMSABERTPPFKITL---MNTAS:QP 660
QY 316 YIDPENV-----KKTGNVTLATGIGSTTANKTVLVDYKGYKFNLSIKG-TID-QIDK 367
Db 661 DISHNVIVPRQAGFINANATIALAKNPSVTVSSNGYPGVELKFKDRLNFQVKF 720
QY 368 TNNYRQTIY--VNPBGDNIAPVLGTNLKPNNTDSNALIDQQNTSIKVKVDNAA--DLSES 425
Db 721 TNRNKALTYKLANNGK--SDVYTS---ATDSSAVLYDK-----KIDGASVKASGD 767
QY 426 YFVNPENEDVTVNSNITFPN---ENQYK---VEENTPD-DQITTPYIVVNGHIDPNK 478
Db 768 IFVPANSTKELT--LTLTLPDFKENQYVEGELTFENSSDSSQLRPLYM-----GFFGDWAS 821
QY 479 GDALRSTL 487
Db 822 SDSLPIFASL 830

QY 365 -----IDKTNNTYRQTIYVNPBGDNIAPVLGTNLKPNNTDSNALIDQQNTSIK 413
Db 2271 NMSGTSMASPFATGAQALVQSMNDKNGAFYATYQKMSABERTPPFKITL---MNTAS:QP 660
QY 414 Y-----KVDNAADLSESYFVNPE-----NFDVTN--SVNITFPNPNQYKVFENTPD 458
Db 2329 LMGKKKDEHCECSACRLTEQELBEIKKRKIELENMTSEQKIIISEPSDRKRSVDFSNPS 2388
QY 459 ---DQITTPYIV--VNGHIDPNKSGDLALRSTLYGYNSNII-----WRSMSDN 503
Db 2389 QISUNVEFSPVPLRIETEHAPAPSTSSPPLSTRIYYLSPKVSETVTTTHOWKDGAPM 2448
QY 504 EVAFNNGSGSDG 516
Db 2449 DEIFSPVSTADG 2461

RESULT 12

US-10-732-923-10252
; Sequence 10252, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10252
; LENGTH: 3167
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-732-923-10252

Query Match 5.9%; Score 159; DB 6; Length 3167;
Best Local Similarity 20.1%; Pred. No. 0.043;
Matches 123; Conservative 97; Mismatches 203; Indels 190; Gaps 30;

QY 12 SKEADASENSVTQSDSASNESKSNDSVSAAPKTD--TNVSDTKTSSN-TNNGETSVA 68
Db 1931 TKQEKQORSVVVETSPSNKHIS-DESSIS---MDEIFSRQDNKSTNFKSGSIPIL 1985
QY 69 ONPAQOQETQSSSTNATTEPTVGEATTT-----TNQA----- 103
Db 1986 VLPGEKEVASASIN--LNGVLEGGKQKSTDYGREKVKQVENSYSYSTRHMDGASIS 2043
QY 104 -----NTPATQSSNT---NAEE-----LVNQTNETTFDNTVSSVNSPQNSTN 146
Db 2044 LDDIFNTSSTQKTETKIDNSQEPQLSKPVLKSSI LDDLFNNTSGI-----EKT 2096
QY 147 AENVSTTQDTS TEATPSNNEAPQSDASNKDVVQAVN-----TSAPRMRAFLAAVAA 201
Db 2097 SEKTTTTTTTTRTET-----TDYVKRSTSLVDRFGYETATP--PAISIAAISF 2144
QY 202 DAPAAAGTDITNQLTNVTGIDSGTTVYP---HQAGYVKLVNGYFVNSAVKGDTFKITVP 258
Db 2145 DQPS-----SSQAFFRSKHQ-----NLSSLTVP 2169
QY 259 KELNUNGVTSTAKVPPIMAGDOVLANGVIDSGNVIYFTDYVNTKDDVKATLTWPA--Y 316
Db 2170 GKWNESMMSNTSTI-----SLDSDFNNSFSKNTSQV-----EPRMRKPLTLFVDNW 2217
QY 317 IDPENVKKTGNVTATGIGSTTANKTVLVDYKGYKF---YNLSIK-----GTIDQ- 364
Db 2218 ID-----NLVSEATNEATKEAPTKPSDNTLNYFRSPTRISQEIKEYWADMIGDIRK 2270
QY 365 -----IDKTNNTYRQTIYVNPBGDNIAPVLGTNLKPNNTDSNALIDQQNTSIK 413
Db 2271 NMSGTSMASPFATGAQALVQSMNDKNGAFYATYQKMSABERTPPFKITL---MNTAS:QP 660
QY 414 Y-----KVDNAADLSESYFVNPE-----NFDVTN--SVNITFPNPNQYKVFENTPD 458


```
Db 2329 LMGKKDDEHCECSACRLTEQLELEIKRKIELENMTSEQIIISBPSDRKSVDFSNPS 2388
QY 459 ----DOITPYIV--VNGHIDPNSGDALRLTLGYNSNII-----WRMSWDN 503
Db 2389 QISLNEVFPVELRTEHAPAPSTSSPPESTRYYILSPKVSETVTTTHQWKDAGIPM 2448
QY 504 EVAFNNGSGSDG 516
Db 2449 DEIFSPVSTADG 2461

RESULT 13
US-10-732-923-13552
; Sequence 13552, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Egerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 13552
; LENGTH: 2964
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-13552

Query Match 5.8%; Score 157.5; DB 6; Length 2964;
Best Local Similarity 17.9%; Pred. No. 0.049;
Matches 102; Conservative 91; Mismatches 241; Indels 137; Gaps 22;

QY 1 VGTLLGGLLSKEADASENSVTQSDASNESKNSDSSVSAAPKDDTIVNSDTKSSNT 60
Db 195 LSNIIIGFPLGNVLLDP PPPPSNSPPISKTSNNLNVSNNNNNNNNNSLSNS 254
QY 61 NNGETS---VAQNPACQETQSS--SNATTEETPVTEATTTTQACATPATQSSNTN 115
Db 255 GNGDESPDFQSNLVKVSRENSGNSLMLHOTSIPNNNSNVVNNNNNNNNNNNN 314
QY 116 AELVNTQTSNETTFNDTNTVSSVNSPNTSNAENVSTTQDTSTEATPSNNEAPQS--- 171
Db 315 N-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 370
QY 172 TDASNKDVNOAVNT-----SAPRMAFSLAAVAADAPAGTDTTNQLTNVTVG 220
Db 371 TSSGNLRLSNTENSLSCKKVAIDLIQNIPLYLLFSLKI---QQLEGEKILFSAENIFC- 426
QY 221 IDSGTTVYPHQAQGVKLVNGFSVPNSAVKGD---TFKI-----TVPKELNLGV 266
Db 427 VD-----KCNF---PPNKALGEIWTNFRIFINSNSNSTIPNSTSTSI 470
QY 267 TSTA-----KVPPIMA-----GQVLANGVIDSGNVIYTFDYVNTKDDVKATLM 313
Db 471 SSPASTQNLISILMTLFGSYNGGSGPNTATLSGGST-----SSSSNLT 518
QY 314 PAYIDENVKKTGNVLTATIGGTTANKTVLVYKYGKYNLSIKGTIDQIDKTNNTYR 373
Db 519 PITTSIHTSNGAILNPINNLSNS-----NNNNNNNNNNNNNNNNNNNNNN-- 572
QY 374 QTIYVNPSGDNVIAPLVLTGNLKNTDS---NALIDQNTSIKV-----YKVDNAADLSES 425
Db 573 -----INSNSNNNSINSNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 628
QY 426 YFVNPENFEDVTSNVIITFPNPQYKVEENTPDDQITPYIVVNGHI-----DPNSKGD 480
Db 629 FKLYCKDFCKI-----IGFQ-----INSHLVKFKDLTKCS 661
QY 481 LAIRSTLYGYSNIIWRMSGWDNEVAFNNGS 511
Db 662 VPTLDTITFAYNS-----KESSFGNTECFPDHS 688
```

```
RESULT 14
US-10-868-381-58
; Sequence 58, Application US/10868381
; GENERAL INFORMATION:
; APPLICANT: Samal, Siba K.
; TITLE OF INVENTION: Avian Pneumovirus Genes, Recombinant Avian Pneumoviruses and Metho
; TITLE OF INVENTION: Making
; FILE REFERENCE: 1797.0530002
; CURRENT APPLICATION NUMBER: US/10/868,381
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: PCT/US03/38123
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US 60/430,301
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 58
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Avian pneumovirus
US-10-868-381-58

Query Match 5.8%; Score 156.5; DB 6; Length 585;
Best Local Similarity 24.7%; Pred. No. 0.008;
Matches 84; Conservative 35; Mismatches 140; Indels 81; Gaps 14;

QY 12 SKEADASENSVTQSDASNESKNSDSSVSA-----APK-----TDDTNVSDTK 55
Db 141 SKQTTEPGTSTQKHTPEAPSSKSPPTTQATAOPTTAPKASTAPKNQATTKKTTETDTT 200
QY 56 TSS---NTNN-GETSVAQNPACQETQSSS--TNATTEETPVTEATTTTQANTPAT 109
Db 201 TASRARNNTNPTETATTTPKATTEGKGEGTQHTKQPET-TARETTPOPRRTASR 259
QY 110 QSNTNAEELVNQTSNETTFNDTNTVSSVNSPNTS-----NAENVSTT----- 153
Db 260 PATTKIEBAETTKRTTKN-TQTSGPRPRFSPKSTATENNKRITTKPNTASTD 318
QY 154 -----QDTSTEATPSNNEAPQ-----STDASN-----KDVVNQAVNTS 187
Db 319 SRQOTRITAEQDQOTQTRAKPTTNGAHPOTTTTPEHNTDTTNTSGKSPKEDKTRDPSSK 378
QY 188 APRMAFSLAAVAADAPAG-----TDITNQLTNVTVGIDSGTTVYPHQAQY 234
Db 379 TPTEQEDASKGTAANPGGSAAEADRRAPPATTTGRTTESAAGTTGDSGAETRRRSAA 438
QY 235 VLKLVGFSVPNSAVKGDTFKITVPKELNLNGVTSTAKVPP 274
Db 439 DRRTSGSTAAEAGTAQSGRAT-PKQPS-GGTAAGNTAPP 476

RESULT 15
US-10-955-952-224
; Sequence 224, Application US/10955952
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 3, 2004, 18:53:31 ; Search time 109.231 Seconds
(without alignments)
2781.239 Million cell updates/sec

Title: US-09-679-643-2_COPY_23_550
Perfect score: 2702
Sequence: 1 VGTLLGFLSSKEADASEN.....NGSGSGDGDKPVWPPQDE 528

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2702	100.0	933	2 Q53653	Q53653 staphylococ
2	2596.5	96.1	928	2 Q6GB45	Q6GB45 staphylococ
3	2596.5	96.1	946	2 Q8XJ1	Q8XJ1 staphylococ
4	2564	94.9	1029	2 Q6G1K4	Q6G1K4 staphylococ
5	2466	91.3	935	2 Q932C5	Q932C5 staphylococ
6	2466	91.3	989	2 Q99VJ4	Q99VJ4 staphylococ
7	1257	46.5	881	2 Q93KH7	Q93KH7 staphylococ
8	520	19.2	961	2 Q99RD3	Q99RD3 staphylococ
9	520	19.2	961	2 Q7A3J8	Q7A3J8 staphylococ
10	515	19.1	677	2 Q6EVM0	Q6EVM0 staphylococ
11	515	19.1	940	2 Q53682	Q53682 staphylococ
12	494.5	18.3	1166	2 Q86489	Q86489 staphylococ
13	494	18.3	943	2 Q8NUU8	Q8NUU8 staphylococ
14	494	18.3	957	2 Q6G6H4	Q6G6H4 staphylococ
15	494	18.3	1141	2 Q99W46	Q99W46 staphylococ
16	494	18.3	1141	2 Q932F7	Q932F7 staphylococ
17	492	18.2	1141	2 Q6GBS4	Q6GBS4 staphylococ
18	492	18.2	1141	2 Q8NXX5	Q8NXX5 staphylococ
19	486	18.0	965	2 Q6GDU5	Q6GDU5 staphylococ
20	481	17.8	1137	2 Q6GJA6	Q6GJA6 staphylococ
21	472	17.5	1056	2 Q8C072	Q8C072 staphylococ
22	472	17.5	1092	2 Q70022	Q70022 staphylococ
23	471.5	17.5	931	2 Q9K113	Q9K113 staphylococ
24	467	17.3	877	2 Q99R07	Q99R07 staphylococ
25	467	17.3	877	2 Q7A382	Q7A382 staphylococ
26	464	17.2	913	2 Q86476	Q86476 staphylococ
27	460.5	17.0	741	2 Q6EVM1	Q6EVM1 staphylococ
28	456.5	16.9	1171	2 Q9KWX6	Q9KWX6 staphylococ
29	452.5	16.7	1018	1 FNBA STAAU	FNBA STAAU staphylococ
30	441.5	16.3	1038	2 Q99RD2	Q99RD2 staphylococ
31	441.5	16.3	1038	2 Q7A3J7	Q7A3J7 staphylococ

RESULT 1
Q53653 PRELIMINARY; PRT; 933 AA.
ID Q53653
AC Q53653;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=94224142; PubMed=8170386;
RA McDewitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrogen receptor of Staphylococcus aureus).";
RL Mol. Microbiol. 11:237-248(1994).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (By similarity).
CC EMBL; Z18852; CAA79304.1; -.
DR PIR; S41539; S41539.
DR PDB; 1N67; X-ray; A=202-560.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 933 AA; 97057 MW; EB51A6DE2FF759F4 CRC64;

Query Match 100.0%; Score 2702; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.8e-115;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VGTLLGFLSSKEADASENSVTQSDSAGNESKNSDSSVSAAPKTDNTNVSPTKTSNT	60
Db	23	VGTLLGFLSSKEADASENSVTQSDSAGNESKNSDSSVSAAPKTDNTNVSPTKTSNT	82
QY	61	NGGETSVAQPAQOETQSSNTATTEETPVGTEATTTTNTQANTATTCSSNTNAELV	120
Db	83	NGGETSVAQPAQOETQSSNTATTEETPVGTEATTTTNTQANTATTCSSNTNAELV	142
QY	121	NOTSNFTTNDNTVSSVNSPQNSTNAEVSTTQDTSSTATPSNNSAPOSTDASNKDVV	180
Db	143	NOTSNFTTNDNTVSSVNSPQNSTNAEVSTTQDTSSTATPSNNSAPOSTDASNKDVV	202
QY	181	NOAVNTSAPMRAPSLAAVAADAPAAAGTDITNLTNTVTGIDSGTTTYPHQAGYKLNKG	240

Q6G644 staphylococ
Q8NU10 staphylococ
Q6G6H3 staphylococ
Q8NUU7 staphylococ
Q6GDH2 staphylococ
Q8KR22 staphylococ
Q86488 staphylococ
Q6GBS5 staphylococ
Q8NXX6 staphylococ
Q99W47 staphylococ
Q7A780 staphylococ
Q99W48 staphylococ
Q7A781 staphylococ
Q6GBS6 staphylococ

```
Db 203 NOAVNTSAPRMRAFSLAAVAADAPAAAGTDIINQLTNVTVGIDSGTTVPYHQAGYVKLVNG 262
QY 241 FSVPSNAVKGTFTKIVPEKLNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
Db 263 FSVPSNAVKGTFTKIVPEKLNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVZATLTPAYIDPENVKKTGNVLTATIGISTTANKTTLVDYKYGKFNLSIKG 360
Db 323 VNTKDDVZATLTPAYIDPENVKKTGNVLTATIGISTTANKTTLVDYKYGKFNLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 420
Db 383 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 442
QY 421 DLSSEYFVNPFEDVNTSVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGD 480
Db 443 DLSSEYFVNPFEDVNTSVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGD 502
QY 481 LALRSTLYGNSNIWRSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 528
Db 503 LALRSTLYGNSNIWRSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 550

RESULT 2
Q6GB45
ID Q6GB45 PRELIMINARY; PRT; 928 AA.
AC Q6GB45;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Clumping factor.
GN ORFNames=SA0752;
OS Staphylococcus aureus subsp. aureus MSSA476.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSSA476;
RA Holden M.T.G., Reil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Fellwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagals K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; BX571857; CAG42526.1; -.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 928 AA; 96417 MW; FBE8570209E9B195 CRC64;
```

```
Query Match 96.1%; Score 2596.5; DB 2; Length 928;
Best Local Similarity 96.2%; Pred. No. 1.2e-110;
Matches 508; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

QY 1 VGTLLIGFLLSKKEADASNSVTQSDASNSKNSDSSVSAAAPKTDITNVSDTKTSNT 60
Db 23 VGTLLIGFLLSKKEADASNSVTQSDASNSKNSDSSVSAAAPKTDITNVSDTKTSNT 82
QY 61 NNGETSAQNPAQOETTQSSNTNATETPTVTGEATTTTNCANTPATTQSSNTNAEELV 120
```

```
Db 83 NNGETSAQNPAQOETTQSSALTNATTEETPTVTGEA-TTATNQNATPATTQSSNTNAEELV 141
QY 121 NOTSNETTNDTNTVSSVNSPQNSNAENSVTTQDTSFEATPSNNESAPQSDASNKDQV 180
Db 142 NOTSNETTNDTNTVSSVNSPQNSNAENSVTTQDTSFEATPSNNESAPQSDASNKDQV 201
QY 181 NOAVNTSAPRMRAFSLAAVAADAPAAAGTDIINQLTNVTVGIDSGTTVPYHQAGYVKLVNG 240
Db 202 NOAVNTSAPRMRAFSLAAVAADAPAAAGTDIINQLTNVTVGIDSGTTVPYHQAGYVKLVNG 261
QY 241 FSVPSNAVKGTFTKIVPEKLNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
Db 262 FSVPSNAVKGTFTKIVPEKLNAGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 321
QY 301 VNTKDDVZATLTPAYIDPENVKKTGNVLTATIGISTTANKTTLVDYKYGKFNLSIKG 360
Db 322 VNTKDDVZATLTPAYIDPENVKKTGNVLTATIGISTTANKTTLVDYKYGKFNLSIKG 381
QY 361 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 420
Db 382 TIDQIDKTNNTYRQTIYVNPISGDNVIAPLVLTGNLKNPNTDSNALIDQQNTSIKVKVDNAA 441
QY 421 DLSSEYFVNPFEDVNTSVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGD 480
Db 442 DLSSEYFVNPFEDVNTSVNITFPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGD 501
QY 481 LALRSTLYGNSNIWRSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 528
Db 502 LALRSTLYGNSNIWRSMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 549

RESULT 3
Q8NXJ1
ID Q8NXJ1 PRELIMINARY; PRT; 946 AA.
AC Q8NXJ1;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Name=clfa; OrderedLocNames=MW0764;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yanamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AP004824; BA94629.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptide glycan-anchor.
SQ SEQUENCE 946 AA; 98237 MW; EPEB838793201173 CRC64;
```

```
Query Match 96.1%; Score 2596.5; DB 2; Length 946;
Best Local Similarity 96.2%; Pred. No. 1.2e-110;
Matches 508; Conservative 8; Mismatches 11; Indels 1; Gaps 1;
```

QY 1 VGTLLIGFLLSKEADASENSVTQSDASNEKSDSSVSAAPKTDITDNTVSDTKSSNT 60
DB 23 VGTLLIGFLLSKEADASENSVTQSDASNEKSDSSVSAAPKTDITDNTVSDTKSSNT 82
QY 61 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 120
DB 83 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 141
QY 121 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATPSNNSAPQSDTASNDKVV 180
DB 142 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATPSNNSAPQSDTASNDKVV 201
QY 181 NOAVNTSAPRMARAFSLAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 240
DB 202 NOAVNTSAPRMARAFSLAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 261
QY 241 FSVPSNAVGKDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
DB 262 FSVPSNAVGKDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 321
QY 301 VNTKDDVKATLTPAVIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKG 360
DB 322 VNTKDDVKATLTPAVIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKG 381
QY 361 TIDQIDKTNNTYRQTIYVNPSPGDNVIAVLTKNLTNTDNLIDQNTSIKYYKVDNAA 420
DB 382 TIDQIDKTNNTYRQTIYVNPSPGDNVIAVLTKNLTNTDNLIDQNTSIKYYKVDNAA 441
QY 421 DLSESYFVNPNEDVNTSNITFPNPNQYKVEFTPDQITTPYIYVWNGHIDPNKSGD 480
DB 442 DLSESYFVNPNEDVNTSNITFPNPNQYKVEFTPDQITTPYIYVWNGHIDPNKSGD 501
QY 481 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 528
DB 502 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 549

RESULT 4

Q6GIK4 PRELIMINARY; PRT; 1029 AA.
ID Q6GIK4
AC Q6GIK4; (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Clumping factor.
GN Name=clfa; ORFNames=SA0842;
OS Staphylococcus aureus subsp. aureus MRSA252.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRSA252;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mays R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR ENBL; BX571856; CAG39851.1; -
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.

DR PROSITE; P50847; GRAM_POS_ANCHORING; 1.
KW Cell wall, peptidoglycan-anchor.
SQ SEQUENCE 1029 AA; 106715 MW; 025EL3C6A4C4F020 CRC64;
Query Match 94.9%; Score 2564; DB 2; Length 1029;
Best Local Similarity 94.1%; Pred. No. 4; le-109;
Matches 497; Conservative 17; Mismatches 14; Indels 0; Gaps 0;
QY 1 VGTLLIGFLLSKEADASENSVTQSDASNEKSDSSVSAAPKTDITDNTVSDTKSSNT 60
DB 23 VGTLLIGFLLSKEADASENSVTQSDASNEKSDSSVSAAPKTDITDNTVSDTKSSNT 82
QY 61 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 120
DB 83 NNGETSAVQAQOQETQSSNATTEETPVTEGATTTTNTQANTPATTOSSNTNAEELV 142
QY 121 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATPSNNSAPQSDTASNDKVV 180
DB 142 NQTSNETTNDNTVSSVNSPQNSNAENVTQDTSSTATPSNNSAPQSDTASNDKVV 202
QY 181 NOAVNTSAPRMARAFSLAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 240
DB 203 NOAVNTSAPRMARAFSLAAVAADAPAAAGTDTITNQLTNVTVGIDSGTTVYVPHQAGYVKLYNG 262
QY 241 FSVPSNAVGKDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 300
DB 263 FSVPSNAVGKDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDY 322
QY 301 VNTKDDVKATLTPAVIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKG 360
DB 323 VNTKDDVKATLTPAVIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKG 382
QY 361 TIDQIDKTNNTYRQTIYVNPSPGDNVIAVLTKNLTNTDNLIDQNTSIKYYKVDNAA 420
DB 383 TIDQIDKTNNTYRQTIYVNPSPGDNVIAVLTKNLTNTDNLIDQNTSIKYYKVDNAA 442
QY 421 DLSESYFVNPNEDVNTSNITFPNPNQYKVEFTPDQITTPYIYVWNGHIDPNKSGD 480
DB 443 DLSESYFVNPNEDVNTSNITFPNPNQYKVEFTPDQITTPYIYVWNGHIDPNKSGD 502
QY 481 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 528
DB 503 LALRSTLYGYSNIIWRSMWDNEVAFNNGSGSGDGDIDKPVVPEQDPE 550

RESULT 5

Q932C5 PRELIMINARY; PRT; 935 AA.
ID Q932C5
AC Q932C5;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Name=fnb; OrderedLocustNames=SAV0811;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizukami-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshida K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).


```

Db 248 ANGVINERTIRYTFDYINNKKDLTAELNLFIDPTTVTKQSGQKVEVTGQNKVSKE 307
QY 343 VLVDY-----EKYGFYNLSIKGTIDQIDKNTNTRQTYVYVPSGDNVIAPIVLTGNLKEN 397
Db 308 FDIKYLGVKDRMG-----VTVNGRIDTLNKEGKFSHFAYVKPNQSLTSVTVTGQVTS 363
QY 398 TDSNALIDQNTSISKVYKVDNAADLSESVFV---NPNFEDVTNSVNIFFPNQYKVEF 454
Db 364 YKQSA-----NNPVTKVYKHIGSDELAESVYAKLDTSKFEDVTEKVNLSYTSNGGYTLNL 419
QY 455 NTPDQITTPYIVVNGHIDPNSKGDALRSTLYGNSNIW-----SMSWDNEVA 506
Db 420 GDLDN--SKDYVYKVEGEYDQNAK--DLNFRHLGSHYKYPYPPYPPVQLTWNNGVA 476
QY 507 FNNSGSGDGIDKPVP 523
Db 477 FYSNAKGDKKPNDP 493

RESULT 9
Q7A3J8
ID Q7A3J8 PRELIMINARY; PRT; 961 AA.
AC Q7A3J8;
DT 05-JUL-2004 (TREMRELrel. 27, Created)
DT 05-JUL-2004 (TREMRELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMRELrel. 27, Last annotation update)
DE FNB protein
GN Name=fnbB; OrderedLocusNames=SA2290;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru K.H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AF003137; BAB43593.1; -.
DR InterPro; IPR004237; Fb_bind.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF02986; Fb_bind; 1.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 961 AA; 106010 MW; 364940F894EA4101 CRC64;

Query Match 19.2%; Score 520; DB 2; Length 961;
Best Local Similarity 26.8%; Pred. No. 8.1e-16;
Matches 149; Conservative 89; Mismatches 201; Indels 118; Gaps 15;

QY 1 VCTLIGFLGSSKEADASE---NSVTQSDSASNESKSSNDSSVSAAPKTDITNVSDTKTS 57
Db 21 LGTMIVGGMQGEKAAAEQNNTTVESGSSATESKA-----SETQTT 63
QY 58 SNTNNGETSAQNPAQOQTTQSSSTNATTEPTVTEATTTTNOANTPATTQSSNTNAE 117
Db 64 TNNVNTIDETQSYATSSTEQPSKSTQVTEEAPTQVQAPKVEEMKS-----110
QY 118 ELVNQTSNETTNDNTVSSVNSPQNSNAENVSTTQDTSSTE---ATPSNNEAPQSDTA 174
QY 175 SNKDVVNQAVNTSAPRM--RAFSLAAVA-----ADAPAGTIDITNLTNTVTVGIDSGT---225
Db 134 -----VSEIKPRMKESTDVTAAVEKEVVEETKATGTDVTK-----VEVEEGSEIV 179

```

```

Db 111 -----QEDLPSEKADKETTGTQVDIAQPSN-----136
QY 175 SNKDVVNQAVNTSAPRM-----AFSLAAVAADAPAAAGTIDITNLTNTVTVGID-----SG 224
Db 137 -----VSEIKPRMKESADVTAVSEKEVAEAKATGTDVTKVEVTESSLEGHNKDS 187
QY 225 TTVPHQAGYVKLANGFVSFVPSNAVKGDTFKITVPRKELNLTNGVTSTAKUPPIMAG--DQVL 282
Db 188 NVNPHNAQRVTLKYKWKFGEGIKAGDYFDFTLSDNVETHGISTLRKVPKISTSEDEKVM 247
QY 283 ANGVIDSNGVIVYFTDYVNTKDVKATLTTPAYIDPENVKKTGNVLTATGIGSTTANKT 342
Db 248 ANGVINERTIRYTFDYINNKKDLTAELNLFIDPTTVTKQSGQKVEVTGQNKVSKE 307
QY 343 VLVDY-----EKYGFYNLSIKGTIDQIDKNTNTRQTYVYVPSGDNVIAPIVLTGNLKEN 397
Db 308 FDIKYLGVKDRMG-----VTVNGRIDTLNKEGKFSHFAYVKPNQSLTSVTVTGQVTS 363
QY 398 TDSNALIDQNTSISKVYKVDNAADLSESVFV---NPNFEDVTNSVNIFFPNQYKVEF 454
Db 364 YKQSA-----NNPVTKVYKHIGSDELAESVYAKLDTSKFEDVTEKVNLSYTSNGGYTLNL 419
QY 455 NTPDQITTPYIVVNGHIDPNSKGDALRSTLYGNSNIW-----SMSWDNEVA 506
Db 420 GDLDN--SKDYVYKVEGEYDQNAK--DLNFRHLGSHYKYPYPPYPPVQLTWNNGVA 476
QY 507 FNNSGSGDGIDKPVP 523
Db 477 FYSNAKGDKKPNDP 493

RESULT 10
Q6EVM0
ID Q6EVM0 PRELIMINARY; PRT; 677 AA.
AC Q6EVM0;
DT 01-OCT-2004 (TREMRELrel. 28, Created)
DT 01-OCT-2004 (TREMRELrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMRELrel. 28, Last annotation update)
DE Fibronectin binding protein B precursor.
GN Name=fnbB;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman D2C;
RA Grundmeier M., Hussain M., Becker P., Sinha B.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ629122; CAF32748.1; -.
KW Signal.
FT SIGNAL 1 35 Potential.
FT CHAIN 35 677 fibronectin binding protein B.
SQ SEQUENCE 677 AA; 74737 MW; 37CAD12EFDC160B1 CRC64;

Query Match 19.1%; Score 515; DB 2; Length 677;
Best Local Similarity 26.8%; Pred. No. 9e-16;
Matches 150; Conservative 95; Mismatches 181; Indels 134; Gaps 18;

QY 1 VCTLIGFLGSSKEADASE---NSVTQSDSASNESKSSNDSSVSAAPKTDITNVSDTKTS 57
Db 21 LGTMIVGGMQGEKAAAEQNNTTVESGSSATESKA-----SETQTT 63
QY 58 SNTNNGETSAQNPAQOQTTQSSSTNATTEPTVTEATTTTNOANTPATTQSSNTNAE 117
Db 64 TNNVNTIDETQSYATSSTEQPSQSTQVTEEAPKTQVQAPKVESTRVLP-----112
QY 118 ELVNQTSNETTNDNTVSSVNSPQNSNAENVSTTQDTSSTE---ATPSNNEAPQSDTA 174
Db 113 -----SEKVAOKETITGTQVDIAQPSN-----133
QY 175 SNKDVVNQAVNTSAPRM--RAFSLAAVA-----ADAPAGTIDITNLTNTVTVGIDSGT---225
Db 134 -----VSEIKPRMKESTDVTAAVEKEVVEETKATGTDVTK-----VEVEEGSEIV 179

```



```
QY 226 -----TVYPHOAGYVYKLVNGESVPSNAVKGDTFKLITVPKELNLNGVTSTAKVPPIMAG 278
Db 180 GHKQDTNVVNPNAERVTLYKKVFGGIGKAGDYDFDTLSDNVETHGISTLRKVPEIKST 239
QY 279 D-QVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGIGST 337
Db 240 DGQVMATGEIIGERKVRVYTFKVEYQEKKDLTAELSLNLFIDPTTQKGNQNVKLGEST 299
QY 338 TANKTVLVGY-----EKYGFYNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAPIVLTG 392
Db 300 TVSKIFNIQYLGVRDNGW-----VTANGRIDTLNKVDGKFSHFAYMKPNQSLSSVTVTG 355
QY 393 NL-----KPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVNPEN---FEDVTVNSVNIIEP 445
Db 356 QVTKGNKPGVN-----NPTVKYKHIGSDLLAESVYAKLDDVSKFEDVTNNMSLDFD 407
QY 446 NPNQYKVEFNTPDQITPTPIVYVNGHIDPNSKGDALRLSTLYGYSNIIWRSMWDNEV 505
Db 408 TNGCYSLNFNNLDQ--SKNYVIKGYGYDSNA--SNLEFQTHLFGYVNYYSNLTWKNV 464
QY 506 AFNNGSGSGDGIDK---PVV 522
Db 465 AFYSNNAQGDGDKLKEPII 484

RESULT 11
Q53682
ID Q53682 PRELIMINARY; PRT; 940 AA.
AC Q53682;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibronectin binding protein B.
GN Names:fnbB;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8325-4;
RX MEDLINE=92111475; PubMed=1837266;
RA Jonsson K., Signas C., Muller H.P., Lindberg M.;
RT "Two different genes encode fibronectin binding proteins in
RT Staphylococcus aureus. The complete nucleotide sequence and
RT characterization of the second gene.";
RL Eur. J. Biochem. 202:1041-1048 (1991).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; X62992; CAA44726.1; -.
DR PIR; S19702; S19702.
DR HSSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0009275; C:cell wall (sensu Gram-positive Bacteria); IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR004237; Fc_bind.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF02986; Fc_bind; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YsIRK signal; 1.
DR TIGRfams; TIGR01167; LPXTG anchor; 1.
DR TIGRfams; TIGR01168; YsIRK signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 940 AA; 103555 MW; E35FBBCA907AE345 CRC64;

Query Match 19.1%; Score 515; DB 2; Length 940;
Best Local Similarity 26.8%; Pred. No. 1.3e-15;
Matches 150; Conservative 95; Mismatches 181; Indels 134; Gaps 18;
```

```
QY 1 VGTLLIGFLLSSKEADASE---NSVTQSDSASNESKSDSSVSAAKPTDITDDTNVSDTKTS 57
Db 21 LGTMIIVGMQEKRAAASEQNNTTVEESGSATESKA-----SETQTT 63
QY 58 SNTNNGETSAVONPAQOETTQSSSTNATTEETPVGTGEATTTTNNQANTPATQSSNTNAE 117
Db 64 TNNVNTIDETQSYATSQTSQSTQVTTTEAPKTVQAPKVEISRVDLP----- 112
QY 118 ELVNTQISNETTFNDTNTVSSVNSPQSNNAENSVSTQDSTST---ATPSNNEAQAQSTDA 174
Db 113 -----SEKVADEKETTGTQVDIAQPSN----- 133
QY 175 SNKDVVNQAVNTSAPRM-RAFSLAAVA-----ADAPAAGTDITNLTNTVTGIDSGT--- 225
Db 134 -----VSEIKPRMKRSTDTVAEAEVVEETKATGTDVINK-----VEVEEGSEIV 179
QY 226 -----TVYPHOAGYVYKLVNGESVPSNAVKGDTFKLITVPKELNLNGVTSTAKVPPIMAG 278
Db 180 GHKQDTNVVNPNAERVTLYKKVFGGIGKAGDYDFDTLSDNVETHGISTLRKVPEIKST 239
QY 279 D-QVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATIGIGST 337
Db 240 DGQVMATGEIIGERKVRVYTFKVEYQEKKDLTAELSLNLFIDPTTQKGNQNVKLGEST 299
QY 338 TANKTVLVGY-----EKYGFYNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAPIVLTG 392
Db 300 TVSKIFNIQYLGVRDNGW-----VTANGRIDTLNKVDGKFSHFAYMKPNQSLSSVTVTG 355
QY 393 NL-----KPNTDSNALIDQNTSIKVKYVDNAADLSSEYFVNPEN---FEDVTVNSVNIIEP 445
Db 356 QVTKGNKPGVN-----NPTVKYKHIGSDLLAESVYAKLDDVSKFEDVTNNMSLDFD 407
QY 446 NPNQYKVEFNTPDQITPTPIVYVNGHIDPNSKGDALRLSTLYGYSNIIWRSMWDNEV 505
Db 408 TNGCYSLNFNNLDQ--SKNYVIKGYGYDSNA--SNLEFQTHLFGYVNYYSNLTWKNV 464
QY 506 AFNNGSGSGDGIDK---PVV 522
Db 465 AFYSNNAQGDGDKLKEPII 484

RESULT 12
ID 086489 PRELIMINARY; PRT; 1166 AA.
AC 086489;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sdr E protein.
DE Name:sdr E;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395 (1998).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DR EMBL; AJ005647; CAA06652.1; -.
DR PIR; T28680; T28680.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF05738; Cna_B; 3.
```


Q6G6H4	PRELIMINARY;	PRT;	957 AA.
ID	Q6G6H4		
AC	Q6G6H4;		
DT	05-JUL-2004 (TREMBLrel. 27, Created)		
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)		
DE	Fibronectin-binding protein.		
GN	Name=fnbB; ORFNames=SA2387;		
OS	Staphylococcus aureus subsp. aureus MSSA476.		
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
NCBI_TaxID=282459;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MSSA476;		
RA	Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,		
RA	Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,		
RA	Raeon N., Bentley S.D., Chillingworth C., Chillingworth T.,		
RA	Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,		
RA	Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,		
RA	James K.D., Lennard N., Line A., Mayes R., Moutle S., Muggall K.,		
RA	Armond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,		
RA	Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,		
RA	Spratt B.G., Parkhill J.;		
RT	"Complete genomes of two clinical Staphylococcus aureus strains:		
RT	evidence for the rapid evolution of virulence and drug resistance.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).		
CC	!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by		
CC	an amide bond [by similarity].		
DR	EMBL; BX571857; CAG44201.1; -		
DR	InterPro; IPR004237; Fb bind.		
DR	InterPro; IPR005877; GPCs YSIRK.		
DR	InterPro; IPR001899; Gram_pos_anchor.		
DR	Pfam; PF02986; Fb_bind; 1.		
DR	Pfam; PF00746; Gram_pos_anchor; 1.		
DR	Pfam; PF04650; YSIRK_signal; 1.		
DR	TIGRFAMS; TIGR01167; LEXTG_anchor; 1.		
DR	TIGRFAMS; TIGR01168; YSIRK_signal; 1.		
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.		
KW	Cell wall; Peptidoglycan-anchor.		
SEQUENCE	957 AA; 105980 MW; 28F8BA7FFDFD3EAF CRC64;		
Query Match	18.3%; Score 494; DB 2; Length 957;		
Best Local Similarity	26.5%; Pred. No. 1.2e-14;		
Matches 149; Conservative	87; Mismatches 194; Indels 132; Gaps 17;		
QY	1 VGTILGGLSSKEADASE--NSVTQSDSASNESKSNDSVSAAPKTDITNVSDTKTS 57		
Db	21 LGTIVVMGQEKAAASQNNVTIESSGSSATERKA-----SETQTT 63		
QY	58 SNTNNGETSVQAQPAQOETTQSSSTNATTEPTVGEATTTTNOANTPATQSSNTNAE 117		
Db	64 TNNVNTIDEQSYSATSTEQPSQSTQVTEAPTVAQPKVETSEVDLP----- 112		
QY	118 ELVNQTSNETTFDNTNVSSVNSPQNSTNAENVSTTQDTSTE--ATPSNNESAQSTDA 174		
Db	113 -----SEKVDKETTGTQVDTAQPSN----- 133		
QY	175 SNKDVVNQAVNTSAPMR-----AFSLAAVAADAPAAAGTDITNQLTNVTVGIDSGT--- 225		
Db	134 -----VSEIKPRKRSRTDVTAVTEKEVEEAKATGTDVTSK-----VEVEEGSEIV 179		
QY	226 -----TVYPHOAGVVKLNYGFSVPNSAVKGYTFKITVPKELNLNGVTSKAVPPIMA 277		
Db	180 GHNNKETNVVNPHNAERVTLKYKWKFDGIGKPGDYFOFTLSNNVETHGISPLRKVPDIKS 239		
QY	278 -GQVLANGVIDSDGNVIYFTDQVYNTKDDVKALTMTPAYIDPENVKKTGNVTLATGIGS 336		
Db	240 KDNILAVGKVMDEKIRYFTDYINNKNLMAELNLFIDFTPTVKQKQTVVEVKLGE 299		
QY	337 TTAKTVLVDY----EKYKGFNLSIKGTIDQIDKTNHYRQTIYVNPSPGDNVIAPVLT 391		
Db	300 NKISKPEFDIKYLGVKONWG-----VTNGRITLLDKNSKIHHLAYINPKKSDMTSITN 355		

QY	392	GNLKPTNSNALIDQNTSIVKIVYVDNAADLSSEYFVN----	PENFEDVTSVN--ITFVN	448
Db	356	GFFA----KGLYTGNGVTVKYYELRSDLPESVYANTNDQEKFDVNDMSDKLTSE	411	
QY	447	ENQYKVFENTDDQITTIYIVVNGHIDPNSKGDIALRSLTYGYSNIWR-----	SMSW	501
Db	412	NGSYKL----TLDLNKKGYVVSFEKYNENDK-ELLFRNLHGHYHANYGYYYYYPVVSLTW	467	
QY	502	DNEVAFNNGSGSGDGIDKXPVVP	523	
Db	468	DNGVAFYSNNAQSGDGKDKENDP	489	

RESULT	15
Q99W46	
ID	Q99W46
AC	Q99W46;
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Set-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE	protein.
GN	Name=sdRE; OrderedLocusNames=SA0521;
OS	Staphylococcus aureus (strain N315);
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX	NCBI_TaxID=158879;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=N315;
RX	MEDLINE=213111952; PubMed=11418146;
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA	Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA	Kanehisa M., Yanashita A., Oshima K., Furuya K., Shiba T.,
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus
RT	aureus.";
RL	Lancet 357:1225-1240 (2001).
CC	-!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC	an amide bond (By similarity).
CC	EMBL; AP003131; BAB41752.1; -.
DR	PIR; E89824; E89824.
DR	GO; GO:0005996; C:cell surface; IEA.
DR	GO; GO:0005618; C:cell wall; IEA.
DR	GO; GO:0016020; C:membrane; IEA.
DR	InterPro; IPR008454; Cna B.
DR	InterPro; IPR005877; GpoS_YSIIRK.
DR	InterPro; IPR001899; Gram_pos_anchor.
DR	Pfam; PF05738; Cna B; 3.
DR	Pfam; PF00746; Gram_pos_anchor; 1.
DR	Pfam; PF04650; YSIIRK signal; 1.
DR	TIGRFAMS; TIGR01167; LPXTG anchor; 1.
DR	TIGRFAMS; TIGR01168; YSIIRK signal; 1.
DR	PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW	Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ	SEQUENCE 1141 AA; 124026 MW; 445419D0B8C5A4F8 CRC64;

```

Query Match      18.3%; Score 494; DB 2; Length 1141;
Best Local Similarity 28.0%; Pred. No. 1.5e-14;
Matches 169; Conservative 92; Mismatches 218; Indels 124; Gaps 26;

Qy   1 VGTLLGFLLSKEADASENSVTQSASNESKSNDSSSVSAAPKTDDTINVSDTKTSNT 60
     |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db   37 VGTTLLFG-LGNQEAKAAENTSTE-NAKQDDAITSQNKEV-----VSETENNSTT 84

Qy   61 NNGETSVQAQPAQOETTQSSTNATTEETPVTGEATTTTTNQANTPATTOSSNTNAR--EL 119
     |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db   85 ENDST---NPIRKEINTDSQPEAKEEST----TSSTQQQNNVTATTETKQPNIKEN 135

Qy   120 VNQTSNETTFNDTNV-----SSVNSPQNSNTNAENVUS--T 152
     |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

Db

136

VKPSDKTATDEDTSVILLEBKAPVNTNDVTTKPSTSEIQTKPTTTPQBSTNIENSQPOPT

195

Qy

153

TQD7STEATPSNNE\$AP-----QSTDA\$NKQDVVNQAVNT\$APR

190

Db

196

PSKVDNQVTDATNPKEPVN\$KSELKNNPEKLKELVRNDNNTDR\$TKFVATAPT\$VAPKR

255

Qy

191

MRA\$SLAAVAADAPAA\$GTDITNQLT-----NVTVGIDSGCTTVYPHQAGVYKLN\$YGF\$V\$PNS

246

Db

256

LNAXMRFAVAQFAAVASNNVNDLIITVTKQTKVGGKDNVAAAHDKDIEYDTEFTIDNK

315

Qy

247

AVKGDTEKI-----TVPKELMLNGVT\$TAKVPPIMAGD---QVL\$ANGVID-SDGNVIYTF

297

Db

316

VKKGDTMTIN\$YDKNVI\$PSDL-----TDKNDPIDITDPSGRVIAKGT\$FDRATKQIITVTF

368

Qy

298

TDYVNTKDDVKATILMPAYIDPENY--KKTGNVT\$LATGIGSTTANKTVLVDY\$EYK\$YKFYN

355

Db

369

TDYV\$DKYEDIKARLTL\$Y\$YIDKQAVENET\$SLMTATAGKETSQN--V\$VDYQDPMVHGD

426

Qy

356

LSIKGTIDQIDKNTNNTYRQTIYVNP---SGDNVIA\$PVL-----TG\$NLPNTDSNALIDQ

406

Db

427

SNIO\$IFTKLDENKQTI\$QOIYVNP\$KKTATNTKVDIAG\$QVDDYGNIKLNG\$S-TIID-

484

Qy

407

QNT\$IKVYKVDNAADL\$BSYFV-NPENFEDV\$N\$VNIITFPNPNQYKVEFNTPD-DOI\$TTP

464

Db

485

QNT\$IKVYKVNPNQQLPQ\$NRIYDF\$QYEDVTSQ---FDNKK\$F\$NNVATILDFGDI\$NSA

540

Qy

465

YIVV\$N\$GHIDP\$N\$K\$GDL-----ALR\$T-LYGN\$N\$IIWR\$MSW\$DNEVAFN\$NG\$SGDGI

517

Db

541

YIIK\$V\$SKYTP\$TSDGELDIAQGT\$MRTTDKYG\$-----YNYAGY\$N\$FIVT\$N\$D\$TGGDGT

595

Qy

518

DKP

520

Db

596

VKP

598

Search completed: November 3, 2004, 19:10:17
Job time : 111.231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:48:10 ; Search time 17.8038 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425

Perfect score: 487

Sequence: 1 TLTPAYIDPENVKKTGNT.....NVIAPVLGTGKLTDSNAL 94

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_23Sep04:*
- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	487	100.0	331	5	AAE29263	Aae29263	Staphyloc
2	487	100.0	345	2	AAW31555	Aaw31555	Fibronect
3	487	100.0	496	5	AAU75490	Aau75490	S. aureus
4	487	100.0	496	6	ADA89664	Ada89664	Staphyloc
5	487	100.0	496	6	ADA89663	Ada89663	Staphyloc
6	487	100.0	520	5	AAE29262	Aae29262	Staphyloc
7	487	100.0	927	6	ABM72221	Abm72221	Staphyloc
8	487	100.0	933	3	AAV58435	Aay58435	Staphyloc
9	487	100.0	933	4	AAE69508	Aab69508	Staphyloc
10	487	100.0	933	6	ABU18947	Abj18947	Pathogen
11	487	100.0	936	2	AAW89801	Aaw89801	Staphyloc
12	422	86.7	935	6	ABU16402	Abu16402	Protein e
13	422	86.7	1021	4	AAU33975	Aau33975	Staphyloc
14	422	86.7	1021	4	AAU36951	Aau36951	Staphyloc
15	113	23.2	1018	1	AAU82115	Aap82115	Fibronect
16	113	23.2	1018	4	AAU37245	Aau37245	Staphyloc
17	113	23.2	1018	4	AAU34301	Aau34301	Staphyloc
18	113	23.2	1018	6	ABJ18922	Abj18922	Pathogen
19	113	23.2	1018	6	ABM72537	Abm72537	Staphyloc
20	113	23.2	1027	2	AAW89806	Aaw89806	Staphyloc
21	102.5	21.0	316	7	ABM79017	Abm79017	Staphyloc
22	102.5	21.0	343	7	ABM79016	Abm79016	Staphyloc
23	102.5	21.0	560	7	ABM79015	Abm79015	Staphyloc
24	102.5	21.0	892	6	ABU42557	Abu42557	Protein e
25	102.5	21.0	892	8	ADO84849	Ado84849	S epiderm

ALIGNMENTS

RESULT 1

AAE29263 ID AAE29263 standard; protein; 331 AA.

XX AAE29263; AC

XX 27-JAN-2003 (first entry)

XX Staphylococcus aureus Clf33 protein.

XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; Clf33; N3 protein; immunological; staphylococcal infection; impetigo; pneumonia; furuncle; septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

XX Staphylococcus aureus.

XX WO200272600-A2.

XX 19-SEP-2002.

XX 28-JAN-2002; 2002WO-US00296.

XX 26-JAN-2001; 2001US-0264072P.

XX 12-MAR-2001; 2001US-0274611P.

XX 18-JUN-2001; 2001US-0298413P.

XX 30-JUL-2001; 2001US-0308116P.

XX (INHI-) INHIBITEX INC.

XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;

XX WPI; 2002-759334/B2.

XX N-PSDB; AAD46862.

XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for treating or preventing Staphylococcus aureus infection e.g. wound infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in a human or animal.

XX Claim 9; Page 70-72; 80pp; English.

XX The invention relates to monoclonal antibody which binds the clumping factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA monoclonal antibody is useful for treating or preventing S. aureus infection in a human or animal, and for inhibiting the binding of staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3 protein is useful for inducing an immunological response in a human or

Abp40469 Staphyloc
Aay83171 Cell wall
Aay70120 Staph. ep
Aaw89804 Staphyloc
Aau34403 Staphyloc
Aau37547 Staphyloc
Aar07070 Fibronect
Abj19111 Pathogen
Ada89470 Staphyloc
Abj18923 Pathogen
Abm72536 Staphyloc
Abm79020 Staphyloc
Aaw41602 Staphyloc
Abm79019 Staphyloc
Aau33960 Staphyloc
Aau37093 Staphyloc
Aay08643 S. aureus
Abj18982 Pathogen
Abu43489 Protein e
Abu15854 Protein e

CC animal. These staphylococcal infections include wound infections, sepsis,
 CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
 CC present sequence is Staphylococcus aureus Clf33 protein

XX SQ Sequence 331 AA;
 Query Match 100.0%; Score 487; DB 5; Length 331;
 Best Local Similarity 100.0%; Pred. No. 4.6e-46;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
 DB 113 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 172
 QY 61 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 94
 DB 173 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 206

RESULT 2
 AAW31555
 ID AAW31555 standard; protein; 345 AA.
 XX AC AAW31555;
 XX DT 27-AUG-2003 (revised)
 XX DT 25-MAR-2003 (revised)
 XX DT 21-MAY-1998 (first entry)
 XX DE Fibronectin-binding MSCRAMM derivative pCF33.
 XX KW Fibronectin; pCF33; collagen binding protein; sepsis; infection;
 XX KW microbial surface component recognising adhesive matrix molecule; MSCRAMM;
 XX KW adhesin; vaccine; immunisation; diagnosis; therapy; mastitis.
 XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers
 XX FT Peptide 1..12
 XX FT /note= "vector pQE30-derived peptide"
 XX WO9743314-A2.
 XX PD 20-NOV-1997.
 XX PF 14-MAY-1997; 97WO-US008210.
 XX PR 16-MAY-1996; 96US-0017678P.
 XX PA (TEXA) UNIV. TEXAS A & M SYSTEM.
 XX PA (UABR-) UAB RES FOUND.
 XX PI Hoeoek M, Patti JM, House-Pompeo K, Sthanam N, Symersky J;
 XX WPI; 1998-008801/01.
 XX PT Antibody that interacts with collagen binding domain of Staphylococcal
 XX cna gene product - useful to prevent bacterial sepsis in animal infected
 XX PT with Staphylococcus aureus.

XX PS Disclosure; Page 91; 143pp; English.
 XX CC This protein comprises Staphylococcus aureus fibronectin-binding
 XX microbial surface component recognising adhesive matrix molecule (MSCRAMM)
 XX derivative pCF33, plus a vector-derived N-terminal peptide. The invention
 XX relates to claimed nucleic acid sequences (see AAR3436-38) encoding S.
 XX aureus collagen binding protein (CBP) epitopes M17, M31 and M55 (see
 XX AAW31552-54) that confer protection against S. aureus infection. CBP
 XX protein and antigenic epitopes are contemplated for use in the treatment
 XX of pathological infections, especially to prevent bacterial adhesion to
 XX collagen. The epitopes are also contemplated for use in the preparation
 XX of vaccines and as carrier proteins in vaccine formulations, as well as
 XX in the formulation of compositions for the prevention of S. aureus

CC infection. pCF33 and pQD (see AAW31556) were used to raise anti-MSCRAMM
 CC polyclonal antibodies used in passive immunisation against bovine
 CC mastitis. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-
 CC AUG-2003 to correct OS field.)

XX SQ Sequence 345 AA;
 Query Match 100.0%; Score 487; DB 2; Length 345;
 Best Local Similarity 100.0%; Pred. No. 4.9e-46;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
 DB 125 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 184
 QY 61 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 94
 DB 185 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 218

RESULT 3
 AAU75490
 ID AAU75490 standard; protein; 496 AA.
 XX AC AAU75490;
 XX DT 23-APR-2002 (first entry)
 XX DE S. aureus antigenic protein associated protein #10.
 XX KW Antigenic protein; vaccine; SREX; antibacterial; antiinflammatory;
 XX KW dermatological; antiulcer; tuberculostatic; immunosuppressive;
 XX KW septicemia; food poisoning; skin disorders; peritonitis; endocarditis;
 XX KW tuberculous; blood infection; sepsis; meningitis; pneumonia;
 XX KW stomach ulcer; gonorrhea; necrotising fasciitis; impetigo;
 XX KW Lyme's disease; gastro-enteritis; dysentery; shigellosis.

XX OS Staphylococcus aureus.
 XX FN WO200198499-A1.
 XX PD 27-DEC-2001.
 XX PF 20-JUN-2001; 2001WO-GB002685.
 XX PR 20-JUN-2000; 2000GB-00014907.
 XX PA (UYSH-) UNIV SHEFFIELD.
 XX PA (BIOS-) BIOSYNEXUS INC.
 XX PI Foster S, McDowell P, Brummell K, Clarke S;
 XX WPI; 2002-106544/14.
 XX PT Identifying antigenic polypeptides expressed by pathogenic organisms
 XX e.g., Staphylococcus aureus during infection, by SREX (serological
 XX identification of antigens by recombinant expression cloning) techniques.
 XX PS Disclosure; Page 72-74; 85pp; English.

XX CC The invention relates to a method for identifying antigenic polypeptides
 XX expressed by pathogenic organisms e.g., Staphylococcus aureus during
 XX infection, by SREX (serological identification of antigens by
 XX recombinant expression cloning) techniques. The method involves providing
 XX a nucleic acid library encoding genes/partial gene sequences of
 XX pathogenic organisms, transforming/transfecting the library into host
 XX cells, contacting the polypeptides expressed by the genes with autologous
 XX antisera derived from an animal infected with, or has been infected with
 XX the pathogen and purifying the nucleic acid encoding the polypeptide or
 XX partial polypeptide binding to the antisera. Also included are the
 XX nucleic acids and polypeptides isolated by the method, vectors and
 XX transformed cells expressing them, a vaccine comprising the polypeptide
 XX and the production of monoclonal antibodies against the polypeptides. The

CC protein and vaccine are useful for immunising an animal (preferably
CC human) against a pathogenic microbe. The proteins and antibodies are
CC useful for manufacturing a medicament for treating Staphylococcus aureus-
CC associated septicemia, food poisoning or skin disorders or
CC Staphylococcus epidermidis-associated septicemia, peritonitis,
CC endocarditis, tuberculosis, blood infections, sepsis, meningitis,
CC pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
CC Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
CC sequence is an S. aureus protein sequence associated with the antigenic
CC proteins of the invention. Note: The present sequence is included in the
CC sequence listing but is not mentioned anywhere else in the specification
XX
SQ Sequence 496 AA;

Query Match 100.0%; Score 487; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 7.8e-46;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTN 60
Db 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTN 391
QY 61 NTRYQTIYVNPNGDNVIAPLVTGNLKPNTDSNAL 94
Db 392 NTRYQTIYVNPNGDNVIAPLVTGNLKPNTDSNAL 425

RESULT 4
ADA89664
ID ADA89664 standard; protein; 496 AA.

AC ADA89664;

XX 20-NOV-2003 (first entry)

XX Staphylococcus aureus antigenic protein #203.

XX antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteraemia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.

XX Staphylococcus aureus

XX WO2003011899-A2.

XX 13-FEB-2003.

XX 02-AUG-2002; 2002WO-GB003606.

XX 02-AUG-2001; 2001GB-00018825.

PR 09-JAN-2002; 2002GB-00000349.

XX (UYSH-) UNIV SHEFFIELD.

PA (BIOS-) BIOSYNEXUS INC.

XX Foster S, Mond J, Clarke S, McDowell P, Brummel K;

XX WPI; 2003-256434/25.

XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
PT useful as a vaccine for immunizing humans against e.g. bacteremia, septic
PT shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or
PT impetigo.

XX Claim 4; Page 161; 189pp; English.

XX The present invention describes an antigenic protein or its part, which

CC is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;
CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, neuroprotective, immunosuppressive,
CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicemia, tuberculosis,
CC bacteria-associated food poisoning, blood infections, peritonitis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, S. aureus-associated septicemia, food-poisoning,
CC skin disorders, S. epidermidis-associated septicemia, peritonitis or
CC endocarditis. The present sequence represents a S. aureus antigenic
CC protein sequence from the present invention.

XX SQ Sequence 496 AA;

Query Match 100.0%; Score 487; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 7.8e-46;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTN 60
Db 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLDYKGYKFNLSIKGTIDQIDKTN 391
QY 61 NTRYQTIYVNPNGDNVIAPLVTGNLKPNTDSNAL 94
Db 392 NTRYQTIYVNPNGDNVIAPLVTGNLKPNTDSNAL 425

RESULT 5

ADA89663

ID ADA89663 standard; protein; 496 AA.

XX ADA89663;

XX 20-NOV-2003 (first entry)

XX Staphylococcus aureus antigenic protein #202.

XX antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteraemia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.

XX Staphylococcus aureus.

XX WO2003011899-A2.

XX 13-FEB-2003.

XX 02-AUG-2002; 2002WO-GB003606.

XX 02-AUG-2001; 2001GB-00018825.

PR 09-JAN-2002; 2002GB-00000349.

PA (UYSH-) UNIV SHEFFIELD.
PA (BIOS-) BIOSYNEXUS INC.
XX
PI Foster S, Mond J, Clarke S, McDowell P, Brummel K;
XX
DR WPI; 2003-256434/25.
XX
XX New antigenic polypeptides from *Staphylococcus aureus* or *S. epidermidis*,
XX useful as a vaccine for immunizing humans against e.g. bacteremia, septic
PT shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or
PT impetigo.
XX
XX Claim 4; Page 160-161; 189pp; English.
PS
XX
XX The present invention describes an antigenic protein or its part, which
CC is for use as a vaccine. The antigenic protein is encoded by an isolated
CC DNA molecule of *Staphylococcus aureus* or *S. epidermidis* partial gene
CC sequences (designated dnaSA and dna SE, respectively; and which encodes a
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
CC composition comprising at least one antigenic protein; (2) a method of
CC immunising an animal against a disease or condition caused by a
CC pathogenic microbe by administering the antigenic protein or the vaccine;
CC (3) an antibody or its binding part obtainable by the method above; (4)
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
CC hybridoma cell line produced by the method of (4); and (6) identifying
CC opsonic antigens expressed by a pathogenic microbe. The antigenic
CC proteins have antibacterial, neuroprotective, immunosuppressive,
CC antiinflammatory, antiulcer, immunostimulant and ophthalmological
CC activities, and can be used in vaccines. The antigenic proteins or
CC vaccines can be used for immunising an animal (specifically a human)
CC against a disease or condition caused by a pathogenic microbe, e.g.
CC bacteraemia, septic shock, organ infection, skin infection, bacterial
CC basal colonisation, bacterial eye infections, septicemia, tuberculosis,
CC bacteria-associated food poisoning, blood infections, peritonitis,
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhea,
CC strep throat, streptococcal-associated toxic shock, necrotising
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
CC dysentery, shigellosis, *S. aureus*-associated septicemia, food-poisoning,
CC skin disorders, *S. epidermidis*-associated septicemia, peritonitis or
CC endocarditis. The present sequence represents a *S. aureus* antigenic
CC protein sequence from the present invention.
XX
SQ Sequence 496 AA;
Query Match 100.0%; Score 487; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 7.8e-46;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 60
Db 332 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 391
QY 61 NTYQTIYVNSGDNVIAPIVLTGNLKPNTDSNAL 94
Db 392 NTYQTIYVNSGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 5
AAE29262
ID AAE29262 standard; protein; 520 AA.
XX
XX AAE29262;
XX
XX 27-JAN-2003 (first entry)
XX
XX *Staphylococcus aureus* Clf40 protein.
XX
XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
XX immunological; *staphylococcal* infection; impetigo; pneumonia; furuncle;
XX septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
XX
XX *Staphylococcus aureus*.
OS
XX
PN WO200272600-A2.
XX
PD 19-SEP-2002.
XX
XX 28-JAN-2002; 2002WO-US002296.
PF
XX
XX 26-JAN-2001; 2001US-0264072P.
PR
XX 12-MAR-2001; 2001US-0274611P.
PR
XX 18-JUN-2001; 2001US-0298413P.
PR
XX 30-JUL-2001; 2001US-0308116P.
XX
XX (INHI-) INHIBITEX INC.
PA
XX
XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
PI
XX
XX WPI; 2002-759834/82.
DR
XX N-PSDB; AAD48661.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
PT treating or preventing *Staphylococcus aureus* infection e.g. wound
PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in
PT a human or animal.
XX
XX Claim 9; Page 67-69; 80pp; English.
PS
XX
XX The invention relates to monoclonal antibody which binds the clumping
CC factor A (ClfA) protein from *Staphylococcus aureus*. The anti-ClfA
CC monoclonal antibody is useful for treating or preventing *S. aureus*
CC infection in a human or animal, and for inhibiting the binding of
CC *staphylococcal* bacteria to fibrinogen or fibrin. The immunogenic fragment
CC *S. aureus* Clf40 protein, *S. aureus* Clf33 protein, or the *S. aureus* N3
CC protein is useful for inducing an immunological response in a human or
CC animal. These *staphylococcal* infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is *Staphylococcus aureus* Clf40 protein
XX
SQ Sequence 520 AA;
Query Match 100.0%; Score 487; DB 5; Length 520;
Best Local Similarity 100.0%; Pred. No. 8.3e-46;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 60
Db 293 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVYDYKYGKFNLSIKGTIDQIDKTN 352
QY 61 NTYQTIYVNSGDNVIAPIVLTGNLKPNTDSNAL 94
Db 353 NTYQTIYVNSGDNVIAPIVLTGNLKPNTDSNAL 386
RESULT 7
ABM72221
ID ABM72221 standard; protein; 927 AA.
XX
XX ABM72221;
XX
XX 20-NOV-2003 (first entry)
XX
XX *Staphylococcus aureus* protein #1461.
DE
XX
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target.
XX
XX *Staphylococcus aureus*.
OS
XX
XX WO200294868-A2.
PN
XX
XX 28-NOV-2002.
PD
XX
XX 27-MAR-2002; 2002WO-IB0002637.
PF
XX
XX 27-MAR-2001; 2001GB-00007661.
XX

XX (CHIR-) CHIRON SPA.
XX Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
XX N-PSDB; ACF73781.
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX Claim 1; SEQ ID NO 2922; 49pp; English.
XX
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus proteins of the invention
XX
XX Sequence 927 AA;
Query Match 100.0%; Score 487; DB 6; Length 927;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 391
QY 61 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 94
DB 392 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 8
ID AAY58435 standard; protein; 933 AA.
AC AAY58435;
XX 27-MAR-2000 (first entry)
XX Staphylococcus aureus fibrinogen binding ClfA protein.
XX ClfA; fibrinogen binding protein; bacterial colonisation;
XX indwelling medical device; staphylococcal infection.
XX Staphylococcus aureus.
XX Key Location/Qualifiers
XX Peptide 1..39
XX /note= "Signal peptide"
XX Region 40..559
XX /note= "Region A"
XX Region 332..550
XX /note= "Fibrinogen-binding region"
XX Region 560..867
XX /note= "Region R"
XX Region 896..900
XX /note= "Gram positive wall-associated consensus motif"
XX US6008341-A.
XX 28-DEC-1999.
XX 22-AUG-1994; 94US-00293728.
XX

PR 22-AUG-1994; 94US-00293728.
XX (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX Foster TJ, Mcdevitt DL;
XX WPI; 2000-096389/08.
XX N-PSDB; AA255832.
XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus
XX aureus, useful for treatment or prevention of infections.
XX Disclosure; Fig 2A-1-4; 35pp; English.
XX
XX This sequence represents the Staphylococcus aureus fibrinogen-binding
XX protein, ClfA. ClfA is an important receptor involved in S. aureus
XX colonisation of indwelling medical devices (e.g., catheters, artificial
XX heart valves). Shortly after implantation, the surfaces of medical
XX devices become coated with host plasma and matrix proteins such as
XX fibrinogen and fibronectin, and there is considerable evidence to suggest
XX that bacterial adherence to fibrinogen/fibrin is important in the
XX initiation of device-related infection. The fibrinogen-binding region of
XX ClfA is thought to reside between residues 332 and 550 in a region
XX designated A. The protein also contains a repeated region (region R)
XX comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus
XX contains features present in surface proteins of other Gram positive
XX bacteria that are responsible for anchoring the protein to the cell wall
XX and cell membrane. ClfA, or its fragments, may be used to block S. aureus
XX colonisation of wounds, to prevent adherence of S. aureus to indwelling
XX medical devices, as vaccines to protect against S. aureus infection
XX (e.g., mastitis in ruminants), to raise specific antibodies, and for
XX diagnosis (by agglutination or immunoassay). The specific antibodies are
XX used for passive immunisation, to block infection of wounds or adhesion
XX of S. aureus and for diagnosis. Nucleotides encoding ClfA and its
XX fragments may be used as diagnostic probes
XX
XX Sequence 933 AA;
Query Match 100.0%; Score 487; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 391
QY 61 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 94
DB 392 NTYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 425
RESULT 9
ID AAB69508 standard; protein; 933 AA.
XX AAB69508;
XX 23-APR-2001 (first entry)
XX Staphylococcus aureus ClfA protein.
XX Staphylococcus aureus; clfA; antibiotic; vaccine;
XX fibrinogen binding protein; bacterial infection; mastitis.
XX Staphylococcus aureus.
XX US6177084-B1.
XX 23-JAN-2001.
XX 19-OCT-1999; 99US-00421868.
XX 22-AUG-1994; 94US-00293728.
XX

OFFICIAL

XX PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX PT Foster TJ, Mcdevitt DL;
XX PI WPI; 2001-181608/18.
XX DR N-PSDB; AAF58593.
XX PR Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
XX PT to prevent infection, promote wound healing, block adherence to
XX PT indwelling medical devices and for diagnosing staphylococcus aureus
XX PT infection.
XX XX
XX PS Claim 5; Fig 2; 30pp; English.
XX CC The present sequence is a novel Staphylococcus aureus fibrinogen binding
XX CC protein. It is useful as a vaccine to protect against human and animal
XX CC infections caused by S.aureus, such as against mastitis, to block
XX CC S.aureus from colonising and infecting a wound, to block adherence of
XX CC S.aureus to indwelling medical devices such as catheters, replacement
XX CC heart valves and cardiac assist devices. The protein can be used to
XX CC diagnose bacterial infections. Polyclonal and monoclonal antibodies
XX CC raised against the fibrinogen binding protein are useful for passively
XX CC immunising against infections caused by S.aureus, to prevent infection of
XX CC a wound and to diagnose bacterial infections
XX SQ Sequence 933 AA;
Query Match 100.0%; Score 487; DB 4; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTANKTVLDVYKGFYNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATGIGSTANKTVLDVYKGFYNLSIKGTIDQIDKTN 391
QY 61 NTRYQTIYVNPFGDNVIAAPVLTGNLKPNTDSNAL 94
DB 392 NTRYQTIYVNPFGDNVIAAPVLTGNLKPNTDSNAL 425
RESULT 10
ID ABJ18947
AC ABJ18947 standard; protein; 933 AA.
XX AC ABJ18947;
XX DT 06-MAR-2003 (first entry)
XX DE Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
XX KW Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
XX KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
XX KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
XX KW autoimmune disease; HIV; hepatitis.
XX OS Staphylococcus sp.
XX FN WO200259148-A2.
XX PD 01-AUG-2002.
XX PF 21-JAN-2002; 2002WO-EP000546.
XX PR 26-JAN-2001; 2001AT-00000130.
XX XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX PA Meinke A, Nagy E, Von Ahlsen U, Klade C, Henicks T, Zauner W;
XX PI Minh DB, Vytvytska O, Etz H, Dyla A, Weichhart T, Hafner M;
XX PI Tempelmaier B;
XX WPI; 2003-075410/07.

XX PT Identifying, isolating and producing hyperimmune serum-reactive antigens
XX PT from a pathogen, for preparing vaccine or medicament for treating or
XX PT preventing e.g. staphylococcal infections, comprises providing antibody
XX PT preparation.
XX PS Example 7; Page 160; 252pp; English.
XX CC The invention relates to a novel method for identifying, isolating and
XX CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX CC allergen, a tissue or host prone to auto-immunity, where the antigens are
XX CC used in a vaccine, comprises providing antibody preparation from a plasma
XX CC pool of a type of animal, or individual sera with antibodies against the
XX CC specific pathogen, tumour, allergen, tissue or host prone to auto-
XX CC immunity. The hyperimmune serum-reactive antigens comprising any of the
XX CC 62 sequences of 53-2261 amino acids fully defined in the specification,
XX CC or their hyperimmune fragments are useful for the manufacture of a
XX CC pharmaceutical preparation, particularly a vaccine against staphylococcal
XX CC infections or colonisation against S. aureus or S. epidermidis. The
XX CC preparation of antibodies is useful for the manufacture of a medicament
XX CC for treating or preventing staphylococcal infections or colonisation
XX CC against S. aureus or S. epidermidis. The antibody preparations may also
XX CC be used for diagnostic and imaging purposes. Other conditions that can be
XX CC treated include cancer, autoimmune diseases or infections caused by viral
XX CC (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX CC sequence represents a staphylococcal protein relating to the method for
XX CC identifying and producing pathogen specific antigens of the invention
XX SQ Sequence 933 AA;
Query Match 100.0%; Score 487; DB 6; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTANKTVLDVYKGFYNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATGIGSTANKTVLDVYKGFYNLSIKGTIDQIDKTN 391
QY 61 NTRYQTIYVNPFGDNVIAAPVLTGNLKPNTDSNAL 94
DB 392 NTRYQTIYVNPFGDNVIAAPVLTGNLKPNTDSNAL 425
RESULT 11
ID AAW89801
XX AC AAW89801 standard; protein; 936 AA.
XX AC AAW89801;
XX DT 16-MAR-1999 (first entry)
XX DE Staphylococcus aureus protein SEQ ID #5249.
XX KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
XX KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
XX KW skin infection; surgical wound infection; scalded skin syndrome;
XX KW toxic shock syndrome.
XX OS Staphylococcus aureus.
XX FN EP786519-A2.
XX PD 30-JUL-1997.
XX PF 07-JAN-1997; 97EP-00100117.
XX PR 05-JAN-1996; 96US-0009861P.
XX XX (HUMA-) HUMAN GENOME SCI INC.
XX PA Kunsch CA, Choi GH, Barash SC, Dallon PJ, Fannon MR, Rosen CA;
XX PI WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
PT stored on computer readable medium and used in the production of anti-
PT *S. aureus* vaccines.
XX
PS Claim 23; Page 3255-3258; 3271pp; English.
XX
CC This sequence represents a *Staphylococcus aureus* protein sequence of the
CC invention. The DNA sequences encoding the *S. aureus* proteins are recorded
CC on a computer readable medium, preferably selected from a floppy or hard
CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
CC Homology searches using the *S. aureus* DNA sequences allows putative
CC functions to be assigned so that protein-encoding or regulatory regions
CC of commercial, therapeutic or industrial importance can be obtained.
CC Specifically, sequences which are likely to encode antigens have been
CC identified and these polypeptides can be used in a vaccine composition
CC against *S. aureus* infection. The polypeptides can also be used in a kit
CC for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated
CC in numerous human diseases, including cellulitis, eyelid infections, food
CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
CC DNA sequences can be used for recombinant production of the polypeptides.
CC The new DNA sequences (and their fragments) are useful as primers or
CC probes for isolating homologues of any of the 5191 *S. aureus* DNA sequences
CC contained on the computer readable medium
XX
SQ Sequence 936 AA;

Query Match 100.0%; Score 487; DB 2; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.8e-45;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVTLAGTGTANKTVLDYKYGKFNLSIKGTIDQIDKTN 60
DB 341 TLTPAYIDPENVKKTGNVTLAGTGTANKTVLDYKYGKFNLSIKGTIDQIDKTN 400

QY 61 NTRYQTIYVNPNGDNVIAPIVLTGNLKENTDSNAL 94
DB 401 NTRYQTIYVNPNGDNVIAPIVLTGNLKENTDSNAL 434

RESULT 12
ABU16402
ID ABU16402 standard; protein; 935 AA.
AC
AC ABU16402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #1929.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX *Staphylococcus aureus*.
XX
XX W0200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002W0-US0009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.

DR N-PSDB; ACA20272.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 44326; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 935 AA;

Query Match 86.7%; Score 422; DB 6; Length 935;
Best Local Similarity 84.9%; Pred. No. 3.6e-38;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTMPAYIDPENVKKTGNVTLAGTGTANKTVLDYKYGKFNLSIKGTIDQIDKTN 61
DB 333 LTMPAYIDPENVKKTGNVTLAGTGTANKTVLDYKYGKFNLSIKGTIDQIDKTN 392

QY 62 TYRQTIYVNPNGDNVIAPIVLTGNLKENTDSNAL 94
DB 393 TYRQTIYVNPNGDNVIAPIVLTGNLKENTDSNAL 425

RESULT 13
AAU33975
ID AAU33975 standard; protein; 1021 AA.
AC
AC AAU33975;
XX
XX 14-FEB-2002 (first entry)
XX
XX *Staphylococcus aureus* cellular proliferation protein #251.
XX
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX
XX *Staphylococcus aureus*.
XX
XX W0200170955-A2.
XX
XX 27-SEP-2001.

```
XX 21-MAR-2001; 2001WO-US009180.
PF 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS51834.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 5471; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1021 AA;
XX
XX Query Match 86.7%; Score 422; DB 4; Length 1021;
XX Best Local Similarity 84.9%; Pred. No. 4e-38;
XX Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
XX
XX 2 LTMAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 61
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 333 ITMPAYIDPENVTGTNLTGIGTNTASKTVLIDYKGYGFNLSIKGTIDQIDKTN 392
XX
XX 62 TYRQTIYVPSGDNVIAFVLGKLPNTDSNAL 94
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 393 TYRQTIYVPSGDNVLPALGKLPNTKSNAL 425
XX
XX RESULT 14
XX AAU36951
XX ID AAU36951 standard; protein; 1021 AA.
XX AC AAU36951;
XX
XX 14-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #1121.
XX
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX
```

```
FN WO200170955-A2.
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS54810.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 12544; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1021 AA;
XX
XX Query Match 86.7%; Score 422; DB 4; Length 1021;
XX Best Local Similarity 84.9%; Pred. No. 4e-38;
XX Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
XX
XX 2 LTMAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 61
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 333 ITMPAYIDPENVTGTNLTGIGTNTASKTVLIDYKGYGFNLSIKGTIDQIDKTN 392
XX
XX 62 TYRQTIYVPSGDNVIAFVLGKLPNTDSNAL 94
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 393 TYRQTIYVPSGDNVLPALGKLPNTKSNAL 425
XX
XX RESULT 15
XX AAP82115
XX ID AAP82115 standard; protein; 1018 AA.
XX AC AAP82115;
XX
XX 25-MAR-2003 (revised)
XX 05-JAN-1990 (first entry)
XX
XX Fibronectin binding protein.
XX
XX Fibronectin binding protein; S. aureus; immunisation; mastitis; ruminants;
XX
```

Search completed: November 3, 2004, 19:03:03
Job time : 18.8038 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 4.60992 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425

Perfect score: 487

Sequence: 1 TLTMPAYIDPENVKGTGNT.....NVIAPVLGNLKPTDSNAL 94

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	100.0	345	3	US-08-856-253-7
2	487	100.0	933	3	US-08-293-728-2
3	487	100.0	933	3	US-09-421-868-2
4	487	100.0	936	4	US-08-956-171B-5249
5	487	100.0	936	4	US-08-781-986A-5249
6	113	23.2	1027	4	US-08-956-171B-5254
7	113	23.2	1027	4	US-08-781-986A-5254
8	102.5	21.0	930	3	US-09-134-001C-5314
9	102.5	21.0	930	4	US-09-386-962C-10
10	100.5	20.6	251	4	US-08-956-171B-5252
11	100.5	20.6	251	4	US-08-781-986A-5252
12	99.5	20.4	582	4	US-09-147-405B-13
13	99.5	20.4	593	4	US-09-147-405B-11
14	99.5	20.4	1092	4	US-09-147-405B-15
15	95	19.5	1132	4	US-09-248-796A-15026
16	88.5	18.2	1166	4	US-09-200-650B-7
17	81.5	16.7	1315	4	US-09-200-650B-5
18	79	16.2	1155	4	US-09-710-279-1780
19	79	16.2	1742	4	US-09-386-962C-4
20	79	16.2	1742	4	US-09-386-959-4
21	73.5	15.1	381	4	US-09-569-611C-47
22	73.5	15.1	400	4	US-09-569-611C-48
23	73.5	15.1	571	6	5200183-17
24	73.5	15.1	640	4	US-09-569-611C-45
25	73.5	15.1	659	4	US-09-569-611C-46
26	73.5	15.1	668	1	US-08-445-050-9
27	73.5	15.1	668	2	US-08-204-691-9

28	73.5	15.1	722	1	US-08-347-718B-1
29	73.5	15.1	722	1	US-08-445-050-3
30	73.5	15.1	722	1	US-08-445-050-7
31	73.5	15.1	722	2	US-08-482-262-1
32	73.5	15.1	722	2	US-08-204-691-3
33	73.5	15.1	722	2	US-08-204-691-7
34	73.5	15.1	722	4	US-09-355-295B-4
35	73.5	15.1	722	6	5200183-3
36	73.5	15.1	742	1	US-08-347-718B-2
37	73.5	15.1	742	2	US-08-482-262-2
38	73.5	15.1	742	6	5200183-2
39	73.5	15.1	745	1	US-08-445-050-2
40	73.5	15.1	745	2	US-08-204-691-2
41	73.5	15.1	745	3	US-08-370-223-13
42	73.5	15.1	745	4	US-09-355-295B-3
43	73.5	15.1	747	3	US-09-347-878-32
44	73	15.0	2780	4	US-10-220-587-2
45	71	14.6	918	4	US-09-200-650B-1

ALIGNMENTS

RESULT 1
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0 Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,253
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,678
FILING DATE: 16-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/SOCKET NUMBER: 33,928
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-856-253-7

Query Match 100.0%; Score 487; DB 3; Length 345;
Best Local Similarity 100.0%; Pred. No. 1e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 60
Db 125 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 184
QY 61 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 94
Db 185 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 218
RESULT 2
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 100.0%; Score 487; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 4.1e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 60
Db 332 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 391
QY 61 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 94
Db 392 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 425

RESULT 3
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 100.0%; Score 487; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 4.1e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 60
Db 332 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 391
QY 61 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 94

Db 392 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 425

RESULT 4

US-08-956-171E-5249
; Sequence 5249, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark S. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5249:

US-08-956-171E-5249
Query Match 100.0%; Score 487; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 4.2e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 60
Db 341 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 400
QY 61 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 94
Db 401 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 434

US-08-956-171E-5249
Query Match 100.0%; Score 487; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 4.2e-51;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 60
Db 341 TLTMPAYIDPENVKKTGNVTATGIGSTTANKTTLVDYKGYKFNLSIKGTIDQIDKTN 400
QY 61 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 94
Db 401 NTYRQTIYVNPSPGNDVIAPIVLTGNLKPNTDSNAL 434

RESULT 5

US-08-781-986A-5249
; Sequence 5249, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch


```
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5254:
SEQUENCE CHARACTERISTICS:
LENGTH: 1027 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5254

Query Match 23.2%; Score 113; DB 4; Length 1027;
Best Local Similarity 25.0%; Pred. No. 6.6e-05;
Matches 23; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 2 LTMPAYIDPENVKKTG--NVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDK 60
DB LEINLFDPKTVQINGNQTITSLNEEQTSKELDVKKDGIGNY-ANLNGSIETFNKAN 368

QY 61 NTYFQTIYVNPSPGNDVIAPVLTGMLKPNIDSN 92
DB NRFSHVAFIKENNGKTTSTVTGTLMKGSNQ 400

RESULT 8
US-09-134-001C-5314
; Sequence 5314, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5314
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match 21.0%; Score 102.5; DB 3; Length 930;
Best Local Similarity 32.9%; Pred. No. 0.0011;
Matches 24; Conservative 13; Mismatches 31; Indels 5; Gaps 2;

QY 2 LTMPAYIDPENVKKTG--NVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDK 58
DB 386 LKLTSYDKSKVPNNNTKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 443

QY 59 TNNYRQTIYVNP 71
DB 444 KNHTVEQTIYINP 456

RESULT 9
US-09-386-962C-10
; Sequence 10, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25

TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5254:
SEQUENCE CHARACTERISTICS:
LENGTH: 930
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-386-962C-10

Query Match 21.0%; Score 102.5; DB 4; Length 930;
Best Local Similarity 32.9%; Pred. No. 0.0011;
Matches 24; Conservative 13; Mismatches 31; Indels 5; Gaps 2;

QY 2 LTMPAYIDPENVKKTG--NVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDK 58
DB 386 LKLTSYDKSKVPNNNTKLDVEYKTLSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 443

QY 59 TNNYRQTIYVNP 71
DB 444 KNHTVEQTIYINP 456

RESULT 10
US-08-956-171E-5252
; Sequence 5252, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5252:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5252:
US-08-956-171E-5252

Query Match 20.6%; Score 100.5; DB 4; Length 251;
```

```

Best Local Similarity 23.2%; Pred. No. 0.00031;
Matches 23; Conservative 25; Mismatches 38; Indels 13; Gaps 3;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGSTANKTANVLVDY-----EKYGFYNLSIKGTIDQI 56
   ||: |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 135 LSLNLFDPITVTQKGNQNVKLGETTTSKIFNIQYILGGVRDNWG---VTANGRIDTL 190
   ||: |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

QY 57 DKTNNTVRQTIYVNPSCDGNVIAPVLITGNL-----KPN TDS 91
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 191 NKVDGKFSHFAYMKPNQSLSSVTVTGQVTKGNKPGVNN 229
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 11
US-08-781-986A-5252
; Sequence 5252, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Scaphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5252:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5252

Query Match 20.6%; Score 100.5; DB 4; Length 251;
Best Local Similarity 23.2%; Pred. No. 0.00031;
Matches 23; Conservative 25; Mismatches 38; Indels 13; Gaps 3;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGSTANKTANVLVDY-----EKYGFYNLSIKGTIDQI 56
   ||: |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 135 LSLNLFDPITVTQKGNQNVKLGETTTSKIFNIQYILGGVRDNWG---VTANGRIDTL 190
   ||: |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

QY 57 DKTNNTVRQTIYVNPSCDGNVIAPVLITGNL-----KPN TDS 91
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 191 NKVDGKFSHFAYMKPNQSLSSVTVTGQVTKGNKPGVNN 229
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 12
US-09-147-405B-13
; Sequence 13, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 19:03:11 ; Search time 1.37768 Seconds
(without alignments)
1078.240 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_425
Perfect score: 487
Sequence: 1 TLTPPAYIDPENVKKTGNVT.....NVIAPVLGTGNLKPNTDSNAL 94

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 66696 seqs, 15802848 residues

Total number of hits satisfying chosen parameters: 66696

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*
1: /cgn2_6/prodata/2/paa/US06 NEW COMB.pap.*
2: /cgn2_6/prodata/2/paa/US07 NEW COMB.pap.*
3: /cgn2_6/prodata/2/paa/US08 NEW COMB.pap.*
4: /cgn2_6/prodata/2/paa/US09 NEW COMB.pap.*
5: /cgn2_6/prodata/2/paa/US10 NEW COMB.pap.*
6: /cgn2_6/prodata/2/paa/US11 NEW COMB.pap.*
7: /cgn2_6/prodata/2/paa/US12 NEW COMB.pap.*
8: /cgn2_6/prodata/2/paa/US60 NEW COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	15.2	94	6	US-10-220-366A-24254
2	67.5	13.9	679	6	Sequence 24254, A
3	64	13.1	1009	6	Sequence 105, App
4	64	13.1	1009	6	Sequence 266, App
5	64	13.1	1009	6	Sequence 272, App
6	64	13.1	1009	6	Sequence 278, App
7	64	13.1	1037	6	Sequence 284, App
8	64	13.1	1037	6	Sequence 13, Appl
9	64	13.1	1037	6	Sequence 19, Appl
10	64	13.1	1037	6	Sequence 25, Appl
11	64	13.1	1037	6	Sequence 28, Appl
12	64	13.1	1037	6	Sequence 36, Appl
13	64	13.1	1037	6	Sequence 39, Appl
14	64	13.1	1037	6	Sequence 264, App
15	64	13.1	1048	6	Sequence 282, App
16	64	13.1	1048	6	Sequence 15, Appl
17	64	13.1	1048	6	Sequence 26, Appl
18	64	13.1	1048	6	Sequence 37, Appl
19	64	13.1	1340	6	Sequence 270, App
20	64	13.1	1340	6	Sequence 17, Appl
21	64	13.1	1340	6	Sequence 27, Appl
22	64	13.1	1340	6	Sequence 38, Appl
23	64	13.1	1340	6	Sequence 276, App
24	63.5	13.0	360	6	Sequence 22722, A
25	62	12.7	1286	6	Sequence 251, App
					Sequence 495, App

ALIGNMENTS

RESULT 1
US-10-220-366A-24254
; Sequence 24254, Application US/10220366A
; GENERAL INFORMATION:
; APPLICANT: HYSEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; CURRENT APPLICATION NUMBER: US/10/220,366A
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 24254
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-366A-24254

Query Match 15.2%; Score 74; DB 6; Length 94;
Best Local Similarity 25.7%; Pred. No. 0.13;
Matches 19; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
QY 14 KKTGNVT-LATIGSTTANKTVLVDEYKYGKFNLSIKGTIDDKNTTYRQTYVNS 72
Db 5 KKKGPNKIRSGKGTDTTNTVTIIRDYGHYTNKLNLEEDVKFLDTYLPFRNQRE 63
QY 73 GDNVIAPVLGTGNLX 86
Db 64 IENMQPITSNEIE 77
RESULT 2
US-10-805-684-105
; Sequence 105, Application US/10805684
; GENERAL INFORMATION:
; APPLICANT: SAKAMOTO, TAKESHI
; APPLICANT: TAKEDA, SHIZU
; TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 58748(70342)
; CURRENT APPLICATION NUMBER: US/10/805,684
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/455,766
; PRIOR FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/459,936
; PRIOR FILING DATE: 2003-04-02

Sequence 4727, Ap
Sequence 13054, A
Sequence 18303, A
Sequence 1047, Ap
Sequence 19147, A
Sequence 19146, A
Sequence 1, Appli
Sequence 19212, A
Sequence 265, App
Sequence 271, App
Sequence 277, App
Sequence 283, App
Sequence 253, App
Sequence 254, App
Sequence 259, App
Sequence 260, App
Sequence 241, App
Sequence 3, Appli
Sequence 20, Appl
Sequence 29, Appl

```

; PRIOR APPLICATION NUMBER: 60/460,103
; PRIOR FILING DATE: 2003-04-02
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 105
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)..(51)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-805-684-105

Query Match          13.9%; Score 67.5; DB 6; Length 679;
Best Local Similarity 35.1%; Pred. No. 7.4; 27; Indels 11; Gaps 4;
Matches 26; Conservative 10; Mismatches 27; Indels 11; Gaps 4;

QY 1 TLTPAYIDPENVKKTGNV-PLATIGS-----TTANKTVLDVYKYGKFNLSIKGTIDQ 55
Db 348 TSTQPAAL-----VSSGSASTLASAINADLPATASADVAADIAYTSKMDAIGKGTWE 403

QY 56 I--DKTNNVTYRQTI 67
Db 404 IYNDLSKNTTGSTI 417

RESULT 3
US-10-837-269-266
; Sequence 266, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 266
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-266

Query Match          13.1%; Score 64; DB 6; Length 1009;
Best Local Similarity 27.8%; Pred. No. 29; 43; Indels 22; Gaps 5;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45
Db 520 LTVVKLDRKEDKYLFTILAKDNGVPLTSNVTVFVSIIDQNDNSPVFTNEY-KFYVP 578

QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPSGDNVIAPVLTKNLKPN 88
Db 579 ENLPRHGTGVLITVDPDYGDNSAVTLSILDENDDFTIDSQTGVIRPN 626

RESULT 4
US-10-837-269-272
; Sequence 272, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 272
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-272

Query Match          13.1%; Score 64; DB 6; Length 1009;
Best Local Similarity 27.8%; Pred. No. 29; 43; Indels 22; Gaps 5;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45
Db 520 LTVVKLDRKEDKYLFTILAKDNGVPLTSNVTVFVSIIDQNDNSPVFTNEY-KFYVP 578

QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPSGDNVIAPVLTKNLKPN 88
Db 579 ENLPRHGTGVLITVDPDYGDNSAVTLSILDENDDFTIDSQTGVIRPN 626

RESULT 5
US-10-837-269-278
; Sequence 278, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 51158-20083.00
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-278

Query Match          13.1%; Score 64; DB 6; Length 1009;
Best Local Similarity 27.8%; Pred. No. 29; 43; Indels 22; Gaps 5;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPAYIDPENVKKTGNVTLA--TGIGSTTANKTVLVD-----YEKYGKFY-- 45

```

```
Db 520 LTVVKKLDREKDKYLFTILAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 578
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPSGDNVIAPIVLTGNLKP 88
Db 579 ENLPRHGTGVLITVTPDYGDN SAVTLSILDENDDFTIDSQTGVIRPN 626

RESULT 6
US-10-837-269-284
; Sequence 284, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 1009
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-284

Query Match 13.1%; Score 64; DB 6; Length 1009;
Best Local Similarity 27.8%; Pred. No. 29;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIAVIDPENVKKTGNVTIA--TGIGSTTANKTVLVD-----YEYKGFY-- 45
Db 520 LTVVKKLDREKDKYLFTILAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 578
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPSGDNVIAPIVLTGNLKP 88
Db 579 ENLPRHGTGVLITVTPDYGDN SAVTLSILDENDDFTIDSQTGVIRPN 626

RESULT 7
US-10-837-269-13
; Sequence 13, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-13

Query Match 13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIAVIDPENVKKTGNVTIA--TGIGSTTANKTVLVD-----YEYKGFY-- 45
Db 520 LTVVKKLDREKDKYLFTILAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 578
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPSGDNVIAPIVLTGNLKP 88
Db 579 ENLPRHGTGVLITVTPDYGDN SAVTLSILDENDDFTIDSQTGVIRPN 626

RESULT 8
US-10-837-269-19
; Sequence 19, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-19

Query Match 13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIAVIDPENVKKTGNVTIA--TGIGSTTANKTVLVD-----YEYKGFY-- 45
Db 543 LTVVKKLDREKDKYLFTILAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 601
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPSGDNVIAPIVLTGNLKP 88
Db 602 ENLPRHGTGVLITVTPDYGDN SAVTLSILDENDDFTIDSQTGVIRPN 649

RESULT 9
US-10-837-269-25
; Sequence 25, Application US/10837269
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Kanner, Steven B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins Entitled
; FILE REFERENCE: 109PID4 Useful in Treatment and Detection of Cancer
; CURRENT APPLICATION NUMBER: US/10/837,269
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: US10/121,024
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/467,002
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-269-25

Query Match 13.1%; Score 64; DB 6; Length 1037;
Best Local Similarity 27.8%; Pred. No. 30;
Matches 30; Conservative 13; Mismatches 43; Indels 22; Gaps 5;

QY 2 LTMPIAVIDPENVKKTGNVTIA--TGIGSTTANKTVLVD-----YEYKGFY-- 45
Db 543 LTVVKKLDREKDKYLFTILAKDNGVPLTNSVTVFVSIIDQNDNSPVFTTHNEY-KFYVP 601
QY 46 -NLSIKGTIDQIDKTNNTYRQ-----TIYVNPSGDNVIAPIVLTGNLKP 88
Db 602 ENLPRHGTGVLITVTPDYGDN SAVTLSILDENDDFTIDSQTGVIRPN 649
```


This Page Blank (uspto)

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	487	100.0	933	2	Q53653	Q53653	staphylococ
2	481	98.8	1029	2	Q6G1K4	Q6G1K4	staphylococ
3	470	96.5	928	2	Q6GB45	Q6GB45	staphylococ
4	470	96.5	946	2	Q8NXJ1	Q8NXJ1	staphylococ
5	422	86.7	935	2	Q932C5	Q932C5	staphylococ
6	422	86.7	989	2	Q93VJ4	Q93VJ4	staphylococ
7	329	67.6	881	2	Q93MH7	Q93MH7	staphylococ
8	113	23.2	741	2	Q8EVM1	Q8EVM1	staphylococ
9	113	23.2	1018	1	FNBA_STAAU	FNBA_STAAU	staphylococ
10	107.5	22.1	965	2	Q6GDU5	Q6GDU5	staphylococ
11	102.5	21.0	931	2	Q9K1I3	Q9K1I3	staphylococ
12	100.5	20.6	677	2	Q6EVM0	Q6EVM0	staphylococ
13	100.5	20.6	940	2	Q53682	Q53682	staphylococ
14	99.5	20.4	1056	2	Q8CQ72	Q8CQ72	staphylococ
15	99.5	20.4	1092	2	Q700R2	Q700R2	staphylococ
16	99.5	20.4	1137	2	Q8GJA6	Q8GJA6	staphylococ
17	99.5	20.4	1171	2	Q9KWX6	Q9KWX6	staphylococ
18	90.5	18.6	943	2	Q8NUH8	Q8NUH8	staphylococ
19	90.5	18.6	957	2	Q6G6H4	Q6G6H4	staphylococ
20	88.5	18.2	1141	2	Q6GBX4	Q6GBX4	staphylococ
21	88.5	18.2	1141	2	Q8NX5	Q8NX5	staphylococ
22	88.5	18.2	1166	2	Q8G4B9	Q8G4B9	staphylococ
23	87.5	18.0	961	2	Q99RD3	Q99RD3	staphylococ
24	87.5	18.0	961	2	Q7A3J8	Q7A3J8	staphylococ
25	87	17.9	1141	2	Q99W46	Q99W46	staphylococ
26	87	17.9	1141	2	Q932F7	Q932F7	staphylococ
27	84.5	17.4	1015	2	Q6G6H3	Q6G6H3	staphylococ
28	84.5	17.4	1015	2	Q8NUTU	Q8NUTU	staphylococ
29	84.5	17.4	1038	2	Q99RD2	Q99RD2	staphylococ
30	84.5	17.4	1038	2	Q7A3J7	Q7A3J7	staphylococ
31	82	16.8	434	2	Q7N3E9	Q7N3E9	photorehab

RT evidence for the rapid evolution of virulence and drug resistance.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 CC -i- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (By similarity).
 DR EMBL; BX571857; CAG42526.1; -.
 DR InterPro; IPR005877; Gpos_Ysirk.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR Pfam; PF04650; Ysirk_signal; 1.
 DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
 DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
 DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
 DR Cell wall; Peptidoglycan-anchor.
 KW SEQUENCE 928 AA; 96417 MW; FBE8570209B9E195 CRC64;
 SQ
 Query Match 96.5%; Score 470; DB 2; Length 928;
 Best Local Similarity 96.8%; Pred. No. 2.8e-37;
 Matches 90; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 LTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIDKTN 61
 DB 332 ITMPAYIDPENVTGTGNVTLTIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIDKTN 391
 QY 62 TYRQTIYVPSGDNVIAPIVLTGNLKPNTDSNAL 94
 DB 392 TYRQTIYVPSGDNVIAPIVLTGNLKPNTDSNAL 424
 RESULT 4
 Q8NXXJ1 PRELIMINARY; PRT; 946 AA.
 ID Q8NXXJ1
 AC Q8NXXJ1; (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Fibrinogen-binding protein.
 GN Name=cIfa; OrderedLocNames=MW0764;
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MW2;
 RX MEDLINE=22040717; PubMed=12043378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA.";
 RL Lancet 359:1819-1827(2002).
 CC -i- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (By similarity).
 DR EMBL; AF004824; BAB94629.1; -.
 DR HSSP; Q3653; IN67.
 DR GO; GO:0009986; C:cell surface; IEA.
 DR GO; GO:0005618; C:cell wall; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR005877; Gpos_Ysirk.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR Pfam; PF04650; Ysirk_signal; 1.
 DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
 DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
 KW Cell wall; Complete proteome; Peptidoglycan-anchor.
 SQ SEQUENCE 946 AA; 98237 MW; EFPB838793201173 CRC64;
 Query Match 96.5%; Score 470; DB 2; Length 946;
 Best Local Similarity 96.8%; Pred. No. 2.9e-37;
 Matches 90; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 LTMPAYIDPENVKKTGNVTATGIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIDKTN 61
 DB 332 ITMPAYIDPENVTGTGNVTLTIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIDKTN 391

DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Clumping factor.
 GN Name=cIfa; ORFNames=SAR0842;
 OS Staphylococcus aureus subsp. aureus MRSA252.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=282458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MRSA252;
 RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
 RA Feltwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
 RA James K.D., Lennard N., Line A., Mayes R., Mungall K.,
 RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
 RA Spratt B.G., Parkhill J.;
 RT "Complete genomes of two clinical Staphylococcus aureus strains:
 RT evidence for the rapid evolution of virulence and drug resistance.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 CC -i- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (By similarity).
 DR EMBL; BX571856; CAG39851.1; -.
 DR InterPro; IPR005877; Gpos_Ysirk.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR Pfam; PF04650; Ysirk_signal; 1.
 DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
 DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
 DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
 DR Cell wall; Peptidoglycan-anchor.
 KW SEQUENCE 1029 AA; 106715 MW; 025E13C6A4C4F020 CRC64;
 SQ
 Query Match 98.8%; Score 481; DB 2; Length 1029;
 Best Local Similarity 98.9%; Pred. No. 2.7e-38;
 Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVTATGIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIDKTN 60
 DB 332 TLTPAYIDPENVTGTGNVTLTIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIDKTN 391
 QY 61 NTYRQTIYVPSGDNVIAPIVLTGNLKPNTDSNAL 94
 DB 392 NTYRQTIYVPSGDNVIAPIVLTGNLKPNTDSNAL 425
 RESULT 3
 Q6GB45 PRELIMINARY; PRT; 928 AA.
 ID Q6GB45
 AC Q6GB45;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Clumping factor.
 GN ORFNames=SAS0752;
 OS Staphylococcus aureus subsp. aureus MSSA476.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=282459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSSA476;
 RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
 RA Feltwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
 RA James K.D., Lennard N., Line A., Mayes R., Mungall K.,
 RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
 RA Spratt B.G., Parkhill J.;
 RT "Complete genomes of two clinical Staphylococcus aureus strains:

```
QY 62 TYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 94
Db 392 TYRQTIYVNPNGDNVIAPIVLTGNLKPNTDSNAL 424

RESULT 5
Q932C5 PRELIMINARY; PRT; 935 AA.
AC Q932C5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein.
GN Name=fnb; OrderedLocusNames=SAV0811;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
CC EMBL; AP003360; BAB56973.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 935 AA; 96949 MW; DC5A2D92CE3BA91C CRC64;

Query Match 86.7%; Score 422; DB 2; Length 935;
Best Local Similarity 84.9%; Pred. NO. 1.5e-32;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 61
Db 333 LTPAYIDPENVTGNTVLTGIGTNTASKTVLVDYKYGKYNLSIKGTIDQIDKTN 392

RESULT 6
Q99VJ4 PRELIMINARY; PRT; 989 AA.
AC Q99VJ4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen-binding protein A, clumping factor.
GN Name=cifa; OrderedLocusNames=SA0742;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
CC EMBL; AP003360; BAB56973.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 935 AA; 96949 MW; DC5A2D92CE3BA91C CRC64;

Query Match 86.7%; Score 422; DB 2; Length 935;
Best Local Similarity 84.9%; Pred. NO. 1.5e-32;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 61
Db 333 LTPAYIDPENVTGNTVLTGIGTNTASKTVLVDYKYGKYNLSIKGTIDQIDKTN 392

RESULT 7
Q93MH7 PRELIMINARY; PRT; 881 AA.
AC Q93MH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Receptor Fbl precursor.
GN Name=fbl;
OS Staphylococcus lugdunensis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=28035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2342;
RX Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
CC EMBL; AF404823; AAK95649.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=N315;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
CC EMBL; AP003311; BAB41975.1; -.
DR PIR; D89852; D89852.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

Query Match 86.7%; Score 422; DB 2; Length 989;
Best Local Similarity 84.9%; Pred. NO. 1.6e-32;
Matches 79; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTVLVDYKYGKYNLSIKGTIDQIDKTN 61
Db 333 LTPAYIDPENVTGNTVLTGIGTNTASKTVLVDYKYGKYNLSIKGTIDQIDKTN 392

RESULT 7
Q93MH7 PRELIMINARY; PRT; 881 AA.
AC Q93MH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Receptor Fbl precursor.
GN Name=fbl;
OS Staphylococcus lugdunensis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=28035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2342;
RX Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (By similarity).
CC EMBL; AF404823; AAK95649.1; -.
DR HSSP; Q53653; IN67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
```

```
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1 39 Potential.
FT CHAIN 40 881 receptin Fbl.
SQ SEQUENCE 881 AA; 94250 MW; D4296C4959C4F19B CRC64;

Query Match 67.6%; Score 329; DB 2; Length 881;
Best Local Similarity 66.7%; Pred. No. 1.8e-23;
Matches 62; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKVLVDYKYGKFNLSIKGTIDQIDKTN 61
Db 317 ISIPGYIDPNVHTGKVNLTSGITQAKTIVVDYKYGKFNLSIKGTIDQIDKVN 376

QY 62 TYRQTIYVNPSSGDNVIAFVLGNLKPNTDSN 94
Db 377 TYRQTIYVNPSSGDNVIAFVLGNLKPNTDSN 409

RESULT 8
Q6EVM1 PRELIMINARY; PRT; 741 AA.
ID Q6EVM1
AC Q6EVM1
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fibronectin binding protein A precursor.
GN Name=fnbA;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman D2C;
RA Grundmeier M., Hussain M., Becker P., Sinha B.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ629121; CAF32747.1; --
KW Signal.
FT SIGNAL 1 36 Potential.
FT CHAIN 37 741 fibronectin binding protein A.
SQ SEQUENCE 741 AA; 81433 MW; 54D9A417128AE7EC CRC64;

Query Match 23.2%; Score 113; DB 2; Length 741;
Best Local Similarity 25.0%; Pred. No. 0.023;
Matches 23; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKVLVDY-EYKGFYNLSIKGTIDQIDKTN 60
Db 301 LEINLFIDPKVTQTNQNTITSLNEEQTSKELDVYKDGIGNY-ANLNGSIETFNKAN 359

QY 61 NTRYQTIYVNPSSGDNVIAFVLGNLKPNTDSN 92
Db 360 NRFSHVAFIKPNNGKTTSTVTGTLMKGSNQ 391

RESULT 9
FNBA_STAAU STANDARD; PRT; 1018 AA.
ID FNBA_STAAU
AC P14738;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Fibronectin-binding protein precursor (FNBP).
GN Name=fnbA;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 8325-4;
RX MEDLINE=89098998; PubMed=25213391;
RA Signaes C., Raucchi G., Joensson K., Lindgren P.-E.,
RA Anantharamaiah G.M., Hoeck M., Lindberg M.;

"Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylococcus aureus: use of this peptide sequence in the synthesis of biologically active peptides.";
Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).
-1- FUNCTION: The ability of bacteria to bind fibronectin has been proposed as a virulence factor enabling bacteria to colonize wound tissues and blood clots. Binding of plasma fibronectin to the bacterial surface might block adhesion receptors on S.aureus, thus representing an important defense mechanism against tissue invasion.
-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; J04151; AAA26632.1; --
InterPro; IPR004237; Fn_bind.
InterPro; IPR005877; GPos_YSIK.
InterPro; IPR001899; Gram_pos_anchor.
Pfam; PF02986; Fn_bind; 1.
Pfam; PF00746; Gram_pos_anchor; 1.
Pfam; PF04650; YSIK_signal; 1.
TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
TIGRFAMs; TIGR01168; YSIK_signal; 1.
PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Cell adhesion; Cell wall; Peptidoglycan-anchor; Repeat; Signal; Virulence.
FT SIGNAL 1 36
FT CHAIN 37 985 Fibronectin-binding protein.
FT PROPEP 986 1018 Removed by sortase (Potential).
FT REPEAT 545 574 B-1.
FT REPEAT 575 604 B-2.
FT DOMAIN 745 878 4 X APPROXIMATE TANDEM REPEATS, FIBRONECTIN-BINDING DOMAIN.
FT REPEAT 745 782 D-1.
FT REPEAT 783 820 D-2.
FT REPEAT 821 859 D-3.
FT REPEAT 860 878 D-4 (incomplete).
FT DOMAIN 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
FT REPEAT 893 906 WR 1.
FT REPEAT 907 920 WR 2.
FT REPEAT 921 934 WR 3.
FT REPEAT 935 948 WR 4.
FT REPEAT 982 986 LPXTG sorting signal (Potential).
FT SITE 985 985 Pentaglycyl murein peptidoglycan amidated threonine (Potential).
FT MOD_RES 985 985
SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020B81F1F CRC64;

Query Match 23.2%; Score 113; DB 1; Length 1018;
Best Local Similarity 25.0%; Pred. No. 0.032;
Matches 23; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 2 LTMPAYIDPENVKKTGNVLTATIGSTTANKVLVDY-EYKGFYNLSIKGTIDQIDKTN 60
Db 301 LEINLFIDPKVTQTNQNTITSLNEEQTSKELDVYKDGIGNY-ANLNGSIETFNKAN 359

QY 61 NTRYQTIYVNPSSGDNVIAFVLGNLKPNTDSN 92
Db 360 NRFSHVAFIKPNNGKTTSTVTGTLMKGSNQ 391

RESULT 10
Q6GDU5 PRELIMINARY; PRT; 965 AA.
ID Q6GDU5
AC Q6GDU5
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
```

05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
Name=fba; ORFNames=SAR2580;
Staphylococcus aureus subsp. aureus MRSA252.
Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=282458;
[1]
SEQUENCE FROM N.A.
RC STRAIN=MRSA252;
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltham T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
EMBL: BX571856; CAG41560.1;
DR InterPro: IPR004237; Fb_bind.
DR InterPro: IPR005877; Gpos_YsIRK.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF02986; Fb_bind; 1.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YsIRK_signal; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YsIRK_signal; 1.
DR PROSITE: PS00847; GRAM_POS_ANCHORING; 1.
DR Cell wall; Peptidoglycan-anchor.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 965 AA; 105691 MW; 2982E132164D0551 CRC64;

Query Match 22.1%; Score 107.5; DB 2; Length 965;
Best Local Similarity 25.0%; Pred. No. 0.11;
Matches 23; Conservative 23; Mismatches 45; Indels 1; Gaps 1;

QY 2 LMPAYIDPENVKKTGNVLTATIGTGSTANKTVLVVDYKYKFNLSIKGTIDQDKTN 61
Db 301 LEINFDIKTQSQNQQTITKNGKSTGTMQITKDGVRNQXTNVNGSIETFDKKN 360

QY 62 TYRQTIYVNP-SGDNVIAPLVTGNLKPNTDSN 92
Db 361 KTHVAYIKPINGSNSDSTVVTGMLTQGSNEN 392

RESULT 11
Q9KIL3
ID Q9KIL3 PRELIMINARY; PRT; 931 AA.
AC Q9KIL3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fibrinogen-binding protein SdrG.
GN Name=sdrG;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).

EMBL: AF245042; AAF72510.1; -
DR HSSP: Q53653; IN67.
DR GO: GO:0009966; C:cell surface; IEA.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR008454; Cna_B.
DR InterPro: IPR005877; Gpos_YsIRK.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF05738; Cna_B; 2.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF04650; YsIRK_signal; 1.
DR TIGRFAMs: TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YsIRK_signal; 1.
DR PROSITE: PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

Query Match 21.0%; Score 102.5; DB 2; Length 931;
Best Local Similarity 32.9%; Pred. No. 0.31;
Matches 24; Conservative 13; Mismatches 31; Indels 5; Gaps 2;

QY 2 LMPAYIDPENVKKTG---NVLTATIGTGSTANKTVLVVDYKYKFNLSIKGTIDQDK 58
Db 387 LKLTSTYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 444

QY 59 TNNVTYRQTIYVNP 71
Db 445 KNHVEQTIYINP 457

RESULT 12
Q6EVM0
ID Q6EVM0 PRELIMINARY; PRT; 677 AA.
AC Q6EVM0;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fibronectin binding protein B precursor.
GN Name=fnbB;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Newman D2C;
RA Grundmeier M., Hussain M., Becker P., Sinha B.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ7629122; CAF32748.1; -
KW Signal.
FT SIGNAL 1 35 Potential.
FT CHAIN 35 677 fibronectin binding protein B.
SQ SEQUENCE 677 AA; 74737 MW; 37CAD12EFD0C160B1 CRC64;

Query Match 20.6%; Score 100.5; DB 2; Length 677;
Best Local Similarity 23.2%; Pred. No. 0.34;
Matches 23; Conservative 25; Mismatches 38; Indels 13; Gaps 3;

QY 2 LMPAYIDPENVKKTGNVLTATIGTGSTANKTVLVVDY-----EYKGFNLSIKGTIDQI 56
Db 273 LSLNLFIDFTTQKGNQNVKLGTTVSKIFNIQYLGVRDNGW----VTANGRIDTL 328

QY 57 DKTNTYRQTIYVNPSPGDNVIAPLVTGNL-----KPTNDS 91
Db 329 NKVDGKFSHFAYMKPNNSQSLSSVTVTGQTKGNKPGVNN 367

RESULT 13
Q53682
ID Q53682 PRELIMINARY; PRT; 940 AA.
AC Q53682;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE DE
GN Name=fnbB;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8325-4;
RX MEDLINE=92111475; PubMed=1837266;
RA Jonsson K., Signas C., Muller H.P., Lindberg M.;
RT "Two different genes encode fibronectin binding proteins in
RT Staphylococcus aureus. The complete nucleotide sequence and
RT characterization of the second gene.";
RL Eur. J. Biochem. 202:1041-1048(1991).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (By similarity).
DE EMBL; X62992; CAA44726.1; -.
DR PIR; S19702; S19702.
DR HSSP; Q53653; 1N67.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0009275; C:cell wall (sensu Gram-positive Bacteria); IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR004237; Fb bind.
DR InterPro; IPR005877; Gpos YSIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF02986; Fb_bind; 2.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSIRK_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_signal; 1.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
CC Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 940 AA; 103555 MW; E35FBBBCA907AE345 CRC64;

Query Match 20.6%; Score 100.5; DB 2; Length 940;
Best Local Similarity 23.2%; Pred. No. 0.5;
Matches 23; Conservative 25; Mismatches 38; Indels 13; Gaps 3;

QY 2 LTWPAYIDPENVKKTGNVTLATGSGTSITANKTVLDVY-----EKVGFYNLSIKGTIDQI 56
Db 273 LSLNLFIDPTTVTKGNQNVKLGITVSKIFNIQYLGGVDRDNWG----VTANGRIDTL 328
QY 57 DKTNNTYRQTIYVNPSPGDNVIAPVLTKNL----KPTDS 91
Db 329 NKVGGKSHFAYMKPNQSLSSVTVTGQTKGNKEGVNN 367

RESULT 14
Q8QC72 PRELIMINARY; PRT; 1056 AA.
ID Q8QC72
AC Q8QC72;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
DE OrderedLocusNames=SE0331;
GN Staphylococcus epidermidis.
OS Bacteria; Firmicutes; Bacillales; Staphylococcus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RT Mol. Microbiol. 49:1577-1593(2003).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by

Query Match 20.4%; Score 99.5; DB 2; Length 1092;
 Best Local Similarity 31.5%; Pred. No. 0.74;
 Matches 23; Conservative 14; Mismatches 31; Indels 5; Gaps 2;
 QY 2 LTMPEAYIDPENVKKTG--NVTLATGIGSTTANKTVLVDVEYKYGKFFYNLSIKGTIDQIDK 58
 Db 389 LKLTSYIDKSKVPNNNTKLDVEYKTAISS--VNKTIIVEYQRPNNENTANLQSMFTNIDT 446
 QY 59 TNNTYRQTIYVNP 71
 Db 447 KNHTVEQTIYINP 459

Search completed: November 3, 2004, 19:10:19
 Job time : 20.4464 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:48:10 ; Search time 41.4791 Seconds
(without alignments)
1894.005 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_550

Perfect score: 1157

Sequence: 1 TLTPAYIDPENVKGTGNTV.....NGSGSGDGDKPVVPEQDPE 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	1157	100.0	331	5	Aae29263 Staphyloc
2	1157	100.0	345	2	Aaw31555 Fibronect
3	1157	100.0	520	5	Aae29262 Staphyloc
4	1157	100.0	927	6	Abm72221 Staphyloc
5	1157	100.0	933	3	Aay58435 Staphyloc
6	1157	100.0	933	4	Aab69508 Staphyloc
7	1157	100.0	933	6	Abj18947 Pathogen
8	1157	100.0	936	2	Aaw89801 Staphyloc
9	991	85.7	935	6	Abul16402 Protein e
10	991	85.7	1021	4	Aau33975 Staphyloc
11	991	85.7	1021	4	Aau36951 Staphyloc
12	863	74.6	496	5	Aau75490 S. aureus
13	863	74.6	496	6	Ada89664 Staphyloc
14	863	74.6	496	6	Ada89663 Staphyloc
15	255	22.0	767	4	Aau34403 Staphyloc
16	255	22.0	767	4	Aau37547 Staphyloc
17	255	22.0	940	2	Aar07070 Fibronect
18	255	22.0	940	6	Abj19111 Pathogen
19	255	22.0	948	6	Ada89470 Staphyloc
20	255	22.0	948	6	Abj18923 Pathogen
21	255	22.0	948	6	Abm72536 Staphyloc
22	244.5	21.1	343	7	Abm79016 Staphyloc
23	244.5	21.1	560	7	Abm79015 Staphyloc
24	244.5	21.1	892	6	Abu42557 Protein e
25	244.5	21.1	892	8	Ado84849 S epiderm

26	244.5	21.1	930	5	ABP40469
27	244.5	21.1	991	3	AY83171 Cell wall
28	244.5	21.1	991	3	AY70120 Staph. ep
29	241.5	20.9	549	7	ABM79020 Staphyloc
30	241.5	20.9	1092	2	AAW41602 Staphyloc
31	241.5	20.9	1092	7	ABM79019 Staphyloc
32	237.5	20.5	961	6	ABU15854 Protein e
33	208	18.0	567	6	ADA89555 Staphyloc
34	208	18.0	743	6	ADA89690 Staphyloc
35	208	18.0	877	6	ADA89539 Staphyloc
36	208	18.0	877	6	ABU42504 Protein e
37	208	18.0	877	6	ABM72702 Staphyloc
38	208	18.0	913	6	ABJ18917 Pathogen
39	208	18.0	918	2	AAy88640 S. aureus
40	207.5	17.9	1166	2	AAy08643 S. aureus
41	207.5	17.9	1166	6	ABJ18982 Pathogen
42	200.5	17.3	1141	6	ABU42327 Protein e
43	196	16.9	316	7	ABM79017 Staphyloc
44	187.5	16.2	670	6	ABU42520 Protein e
45	183.5	15.9	1018	4	Aau37245 Staphyloc

ALIGNMENTS

RESULT 1

AAE29263
ID AAE29263 standard; protein; 331 AA.

XX AC AAE29263;

XX AC (first entry)

DT 27-JAN-2003 (first entry)

XX Staphylococcus aureus Clf33 protein.

XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf33; N3 protein;

KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;

KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.

XX Staphylococcus aureus.

XX WO200272600-A2.

XX 19-SEP-2002.

XX 28-JAN-2002; 2002WO-US002296.

XX 26-JAN-2001; 2001US-0264072P.

XX 12-MAR-2001; 2001US-0274611P.

XX 18-JUN-2001; 2001US-0298413P.

XX 30-JUL-2001; 2001US-0308116P.

XX (INHI-) INHIBITEX INC.

XX Patti JM, Hutchins JT, Dawanski P, Patel P, Hall A;

XX WPI; 2002-759334-02.

XX N-PSDB; AAD46862.

XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for

XX treating or preventing Staphylococcus aureus infection e.g. wound

XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

XX a human or animal.

XX Claim 9; Page 70-72; 80pp; English.

XX The invention relates to monoclonal antibody which binds the clumping

XX factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA

XX monoclonal antibody is useful for treating or preventing S. aureus

XX infection in a human or animal, and for inhibiting the binding of

XX staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment

XX S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3

XX protein is useful for inducing an immunological response in a human or

CC animal. These staphylococcal infections include wound infections, sepsis,
CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
CC present sequence is Staphylococcus aureus Clf33 protein
XX
XX
SQ Sequence 331 AA;

Query Match 100.0%; Score 1157; DB 5; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.le-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 113 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 172

QY 61 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKYKYVDNAADLSSEYFVN 120
DB 173 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKYKYVDNAADLSSEYFVN 232

QY 121 PENFEDVTNSVNIFFPNPNQYKVFETPDQITTPYIVVNGHIDPNSKGDALRSTLYG 180
DB 233 PENFEDVTNSVNIFFPNPNQYKVFETPDQITTPYIVVNGHIDPNSKGDALRSTLYG 292

QY 181 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 219
DB 293 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 331

RESULT 2
AAW31555
ID AAW31555 standard; protein; 345 AA.
XX
XX AAW31555;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 21-MAY-1998 (first entry)
XX
XX
XX Fibronectin-binding MSCRAMM derivative pCF33.
XX
XX Fibronectin; pCF33; collagen binding protein; sepsis; infection;
KW microbial surface component recognising adhesive matrix molecule; MSCRAMM;
KW adhesin; vaccine; immunisation; diagnosis; therapy; mastitis.
XX
XX Staphylococcus aureus.
XX
XX
XX Key Location/Qualifiers
FH Peptide 1..12
FT /note= "vector pQE30-derived peptide"
XX
XX WO9743314-A2.
XX
XX 20-NOV-1997.
XX
XX 14-MAY-1997; 97WO-US008210.
XX
XX 16-MAY-1996; 96US-0017678P.
XX (TEXA) UNIV TEXAS A & M SYSTEM.
XX (UABR-) UAB-RES FOUND.
XX
XX Hoeek M, Patti JM, House-Pompeo K, Sthanam N, Symersky J;
XX WPI; 1998-008601/01.
XX
XX Antibody that interacts with collagen binding domain of Staphylococcal
PT cna gene product - useful to prevent bacterial sepsis in animal infected
PT with Staphylococcus aureus.
XX
XX Disclosure; Page 91; 143pp; English.
XX
XX This protein comprises Staphylococcus aureus fibronectin-binding
CC microbial surface component recognising adhesive matrix molecule (MSCRAMM)
CC derivative pCF33, plus a vector-derived N-terminal peptide. The invention

CC relates to claimed nucleic acid sequences (see AAT93436-38) encoding S.
CC aureus collagen binding protein (CBP) epitopes M17, M31 and M55 (see
CC AAW31552-54) that confer protection against S. aureus infection. CBP
CC protein and antigenic epitopes are contemplated for use in the treatment
CC of pathological infections, especially to prevent bacterial adhesion to
CC collagen. The epitopes are also contemplated for use in the preparation
CC of vaccines and as carrier proteins for the prevention of S. aureus
CC infection. pCF33 and pQD (see AAW31556) were used to raise anti-MSCRAMM
CC polyclonal antibodies used in passive immunisation against bovine
CC mastitis. (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1157; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.le-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 125 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 184

QY 61 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKYKYVDNAADLSSEYFVN 120
DB 185 NTRYQTIVNPSGDNVIAPVLTGNLKTNTSNALIDQONTSIKYKYVDNAADLSSEYFVN 244

QY 121 PENFEDVTNSVNIFFPNPNQYKVFETPDQITTPYIVVNGHIDPNSKGDALRSTLYG 180
DB 245 PENFEDVTNSVNIFFPNPNQYKVFETPDQITTPYIVVNGHIDPNSKGDALRSTLYG 304

QY 181 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 219
DB 305 YNSNIIWRSMWDNEVAFNNGSGGSGDGIDKVPVPEQDPE 343

RESULT 3
AAE29262
ID AAE29262 standard; protein; 520 AA.
XX
XX AAE29262;
XX
XX 27-JAN-2003 (first entry)
XX
XX Staphylococcus aureus Clf40 protein.
XX
XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;
KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;
KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.
XX
XX Staphylococcus aureus.
XX
XX WO200272600-A2.
XX
XX 19-SEP-2002.
XX
XX 28-JAN-2002; 2002WO-US002296.
XX
XX 26-JAN-2001; 2001US-0264072P.
PR 12-MAR-2001; 2001US-0274611P.
PR 18-JUN-2001; 2001US-0298413P.
PR 30-JUL-2001; 2001US-0308116P.
XX
XX (INHI-) INHIBITEX INC.
XX
XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;
PI WPI; 2002-759834/82.
XX
XX N-PSDB; AAD46861.
XX
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for
PT treating or preventing Staphylococcus aureus infection e.g. wound
PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in

PT a human or animal.
 XX Claim 9; Page 67-69; 80pp; English.
 XX The invention relates to monoclonal antibody which binds the clumping
 CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA
 CC monoclonal antibody is useful for treating or preventing S. aureus
 CC infection in a human or animal, and for inhibiting the binding of
 CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment
 CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3
 CC protein is useful for inducing an immunological response in a human or
 CC animal. These staphylococcal infections include wound infections, sepsis,
 CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The
 CC present sequence is Staphylococcus aureus Clf40 protein
 XX
 SQ Sequence 520 AA;

Query Match 100.0%; Score 1157; DB 5; Length 520;
 Best Local Similarity 100.0%; Pred. No. 1.9e-86;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTMPAVIDPENVKKTGNVLTATIGSTTANKTVLVVDYKGYKFNLSIKGTIDIDKTN 60
 Db 293 TLTMPAVIDPENVKKTGNVLTATIGSTTANKTVLVVDYKGYKFNLSIKGTIDIDKTN 352
 QY 61 NTRQTIYVNPSSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 120
 Db 353 NTRQTIYVNPSSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 412
 QY 121 PENFEDVTNSVNTFFPNQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 180
 Db 413 PENFEDVTNSVNTFFPNQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 472
 QY 181 YNSNIIWRSMWNEVAFNNGSGGIDKPKVVPPEQDE 219
 Db 473 YNSNIIWRSMWNEVAFNNGSGGIDKPKVVPPEQDE 511

RESULT 4
 ABM72221
 ID ABM72221 standard; protein; 927 AA.
 AC AC
 XX ABM72221;
 XX 20-NOV-2003 (first entry)
 XX Staphylococcus aureus protein #1461.
 XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.
 XX Staphylococcus aureus.
 OS WO200294868-A2.
 PN 28-NOV-2002.
 PD 27-MAR-2002; 2002WO-IB002637.
 PF 27-MAR-2001; 2001GB-00007661.
 PR (CHIR-) CHIRON SPA.
 XX Masignani V, Mora M, Scarselli M;
 PI WPI: 2003-120786/11.
 DR N-PSDB; ACF73781.
 XX New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.
 XX Claim 1; SEQ ID NO 2922; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus proteins of the invention
 XX
 SQ Sequence 927 AA;

Query Match 100.0%; Score 1157; DB 6; Length 927;
 Best Local Similarity 100.0%; Pred. No. 4e-86;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTMPAVIDPENVKKTGNVLTATIGSTTANKTVLVVDYKGYKFNLSIKGTIDIDKTN 60
 Db 332 TLTMPAVIDPENVKKTGNVLTATIGSTTANKTVLVVDYKGYKFNLSIKGTIDIDKTN 391
 QY 61 NTRQTIYVNPSSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 120
 Db 392 NTRQTIYVNPSSGDNVIAPIVLTGNLKPTDSNALIDQNTSIKVKVDNAADLSESYFVN 451
 QY 121 PENFEDVTNSVNTFFPNQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 180
 Db 452 PENFEDVTNSVNTFFPNQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYG 511
 QY 181 YNSNIIWRSMWNEVAFNNGSGGIDKPKVVPPEQDE 219
 Db 512 YNSNIIWRSMWNEVAFNNGSGGIDKPKVVPPEQDE 550

RESULT 5
 AAY58435
 ID AAY58435 standard; protein; 933 AA.
 AC AAY58435;
 XX 27-MAR-2000 (first entry)
 XX Staphylococcus aureus fibrinogen binding ClfA protein.
 DE ClfA; fibrinogen binding protein; bacterial colonisation;
 XX indwelling medical device; staphylococcal infection.
 KW Staphylococcus aureus.
 OS
 XX Key Location/Qualifiers
 FT Peptide 1..39
 FT /note= "Signal peptide"
 FT Region 40..559
 FT /note= "Region A"
 FT Region 332..550
 FT /note= "Fibrinogen-binding region"
 FT Region 560..867
 FT /note= "Region R"
 FT Region 896..900
 FT /note= "Gram positive wall-associated consensus motif"
 XX
 PN US6008341-A.
 XX 28-DEC-1999.
 PD 22-AUG-1994; 94US-00293728.
 PF 22-AUG-1994; 94US-00293728.
 PR (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX Foster TJ, McDevitt DL;

copy print

PF 19-OCT-1999; 99US-00421868.
XX
XX
PR 22-AUG-1994; 94US-00293728.
XX
XX
FA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
XX
PI Foster TJ, Mcdevitt DL;
XX
XX
XX WPI: 2001-181608/18.
DR N-PSDB; AAF58593.
XX
XX Novel Staphylococcus aureus fibrinogen binding protein useful as vaccine
PT to prevent infection, promote wound healing, block adherence to
PT indwelling medical devices and for diagnosing staphylococcus aureus
PT infection.
XX
XX Claim 5; Fig 2; 30pp; English.
XX
XX The present sequence is a novel Staphylococcus aureus fibrinogen binding
CC protein. It is useful as a vaccine to protect against human and animal
CC infections caused by S.aureus, such as against mastitis, to block
CC S.aureus from colonising and infecting a wound, to block adherence of
CC S.aureus to indwelling medical devices such as catheters, replacement
CC heart valves and cardiac assist devices. The protein can be used to
CC diagnose bacterial infections. Polyclonal and monoclonal antibodies
CC raised against the fibrinogen binding protein are useful for passively
CC immunising against infections caused by S.aureus, to prevent infection of
CC a wound and to diagnose bacterial infections
XX
XX Sequence 933 AA;
SQ
Query Match 100.0%; Score 1157; DB 4; Length 933;
Best Local Similarity 100.0%; Pred. No. 4e-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENKKTGNVTATGIGSTANKTVLDVYKGYKFNLSIKGFIQIDKTN 60
Db 332 TLTMPAYIDPENKKTGNVTATGIGSTANKTVLDVYKGYKFNLSIKGFIQIDKTN 391
QY 61 NYRTQTIYVPSGDNVIAFLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFN 120
Db 392 NYRTQTIYVPSGDNVIAFLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFN 451
QY 121 PENFEDVTSNVTITPENPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 180
Db 452 PENFEDVTSNVTITPENPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 511
QY 181 YNSNIWRSMSWDNEVAFNNGSGGIDKPVVPEQDE 219
Db 512 YNSNIWRSMSWDNEVAFNNGSGGIDKPVVPEQDE 550
RESULT 7
ABJ18947
ID ABJ18947 standard; protein; 933 AA.
XX
XX AC ABJ18947;
XX
XX DT 06-MAR-2003 (first entry)
XX
XX DE Pathogen specific antigen related staphylococcal protein SEQ ID No 93.
XX
XX KW Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
XX hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
XX auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
XX autoimmune disease; HIV; hepatitis.
XX
XX OS Staphylococcus sp.
XX
XX PN WO200259148-A2.
XX
XX PD 01-AUG-2002.
XX

XX WPI: 2000-096389/08.
DR N-PSDB; AAZ55832.
XX
XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus
PT aureus, useful for treatment or prevention of infections.
XX
XX Disclosure; Fig 2A-1-4; 35pp; English.
XX
XX This sequence represents the Staphylococcus aureus fibrinogen-binding
CC protein. ClfA is an important receptor involved in S. aureus
CC colonisation of indwelling medical devices (e.g., catheters, artificial
CC heart valves). Shortly after implantation, the surfaces of medical
CC devices become coated with host plasma and matrix proteins such as
CC fibrinogen and fibronectin, and there is considerable evidence to suggest
CC that bacterial adherence to fibrinogen/fibrin is important in the
CC initiation of device-related infection. The fibrinogen-binding region of
CC ClfA is thought to reside between residues 332 and 550 in a region
CC comprising 154 repeats of the dipeptide Ser-Asp, and the C-terminus
CC contains features present in surface proteins of other Gram positive
CC bacteria that are responsible for anchoring the protein to the cell wall
CC and cell membrane. ClfA, or its fragments, may be used to block S. aureus
CC colonisation of wounds, to prevent adherence of S. aureus to indwelling
CC medical devices, as vaccines to protect against S. aureus infection
CC (e.g., mastitis in ruminants), to raise specific antibodies, and for
CC diagnosis (by agglutination or immunoassay). The specific antibodies are
CC used for passive immunisation, to block infection of wounds or adhesion
CC of S. aureus and for diagnosis. Nucleotides encoding ClfA and its
XX fragments may be used as diagnostic probes
XX
XX Sequence 933 AA;
SQ
Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 4e-86;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENKKTGNVTATGIGSTANKTVLDVYKGYKFNLSIKGFIQIDKTN 60
Db 332 TLTMPAYIDPENKKTGNVTATGIGSTANKTVLDVYKGYKFNLSIKGFIQIDKTN 391
QY 61 NYRTQTIYVPSGDNVIAFLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFN 120
Db 392 NYRTQTIYVPSGDNVIAFLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSSEYFN 451
QY 121 PENFEDVTSNVTITPENPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 180
Db 452 PENFEDVTSNVTITPENPNQYKVEFNTDDQITTPYIVVNGHIDPNSKGDIALRSTLYG 511
QY 181 YNSNIWRSMSWDNEVAFNNGSGGIDKPVVPEQDE 219
Db 512 YNSNIWRSMSWDNEVAFNNGSGGIDKPVVPEQDE 550
RESULT 6
AAB69508
ID AAB69508 standard; protein; 933 AA.
XX
XX AC AAB69508;
XX
XX DT 23-APR-2001 (first entry)
XX
XX DE Staphylococcus aureus ClfA protein.
XX
XX KW Staphylococcus aureus; clfA; antibiotic; vaccine;
XX fibrinogen binding protein; bacterial infection; mastitis.
XX
XX OS Staphylococcus aureus.
XX
XX PN US6177084-B1.
XX
XX PD 23-JAN-2001.
XX

PF 21-JAN-2002; 2002WC-EP000546.

PR 26-JAN-2001; 2001AT-00000130.

PA (CISTEM) CISTEM BIOTECHNOLOGIES GMBH.

PI Meinke A, Nagy E, Von Ahlsen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;
XX Tempelmaier B;

XX WPI; 2003-075410/07.

XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.

XX Example 7; Page 160; 252pp; English.

XX The invention relates to a novel method for identifying, isolating and
XX producing hyperimmune serum-reactive antigens from a pathogen, tumour,
XX allergen, a tissue or host prone to auto-immunity, where the antigens are
XX used in a vaccine, comprises providing antibody preparation from a plasma
XX pool of a type of animal, or individual sera with antibodies against the
XX specific pathogen, tumour, allergen, tissue or host prone to auto-
XX immunity. The hyperimmune serum-reactive antigens comprising any of the
XX 62 sequences of 53-2261 amino acids fully defined in the specification,
XX or their hyperimmune fragments are useful for the manufacture of a
XX pharmaceutical preparation, particularly a vaccine against staphylococcal
XX infections or colonisation against *S. aureus* or *S. epidermidis*. The
XX preparation of antibodies is useful for the manufacture of a medicament
XX for treating or preventing staphylococcal infections or colonisation
XX against *S. aureus* or *S. epidermidis*. The antibody preparations may also
XX be used for diagnostic and imaging purposes. Other conditions that can be
XX treated include cancer, autoimmune diseases or infections caused by viral
XX (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
XX sequence represents a staphylococcal protein relating to the method for
XX identifying and producing pathogen specific antigens of the invention

XX SQ Sequence 933 AA;

Query Match 100.0%; Score 1157; DB 6; Length 933;

Best Local Similarity 100.0%; Pred. No. 4e-86;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 60

DB 332 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 391

QY 61 NTYRQIYVNPDSGDNVIAVLGNLKPNTDSNALIDQNTSIKVKYVDNAADLSESYFVN 120

DB 392 NTYRQIYVNPDSGDNVIAVLGNLKPNTDSNALIDQNTSIKVKYVDNAADLSESYFVN 451

QY 121 PENFEDVTSNLTFFPNQYKVEENTPDQITTPYIVVNGHIDPNSGDLALRSTLYG 180

DB 452 PENFEDVTSNLTFFPNQYKVEENTPDQITTPYIVVNGHIDPNSGDLALRSTLYG 511

QY 181 YNSNIIRWSMDSWNEVAFNNGSGSGDGIDKPVVPEQDE 219

DB 512 YNSNIIRWSMDSWNEVAFNNGSGSGDGIDKPVVPEQDE 550

XX RESULT 8

XX ID AAW89801

XX AC AAW89801;

XX 16-MAR-1999 (first entry)

XX Staphylococcus aureus protein SEQ ID #5249.

XX Computer readable medium; vaccine; *S. aureus* infection; immunodetection;

XX KW

KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
XX toxic shock syndrome.

OS Staphylococcus aureus.

XX EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35

XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -

XX stored on computer readable medium and used in the production of anti-

XX *S. aureus* vaccines.

XX Claim 23; Page 3255-3258; 3271pp; English.

XX This sequence represents a *Staphylococcus aureus* protein sequence of the

XX invention. The DNA sequences encoding the *S. aureus* proteins are recorded

XX on a computer readable medium, preferably selected from a floppy or hard

XX disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.

XX Homology searches using the *S. aureus* DNA sequences allows putative

XX functions to be assigned so that protein-encoding or regulatory regions

XX of commercial, therapeutic or industrial importance can be obtained.

XX Specifically, sequences which are likely to encode antigens have been

XX identified and these polypeptides can be used in a vaccine composition

XX against *S. aureus* infection. The polypeptides can also be used in a kit

XX for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated

XX in numerous human diseases, including cellulitis, eyelid infections, food

XX poisoning, osteomyelitis, skin and surgical wound infections, scalded

XX skin syndrome, toxic shock syndrome, etc. Organisms transformed with the

XX DNA sequences can be used for recombinant production of the polypeptides.

XX The new DNA sequences (and their fragments) are useful as primers or

XX probes for isolating homologues of any of the 5191 *S. aureus* DNA sequences

XX contained on the computer readable medium

XX SQ Sequence 936 AA;

Query Match 100.0%; Score 1157; DB 2; Length 936;

Best Local Similarity 100.0%; Pred. No. 4e-86;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 60

DB 341 TLTPAYIDPENVKKTGNVLTATIGSTTANKTVLDYKGYFYNLSIKGTIDQIDKTN 400

QY 61 NTYRQIYVNPDSGDNVIAVLGNLKPNTDSNALIDQNTSIKVKYVDNAADLSESYFVN 120

DB 401 NTYRQIYVNPDSGDNVIAVLGNLKPNTDSNALIDQNTSIKVKYVDNAADLSESYFVN 460

QY 121 PENFEDVTSNLTFFPNQYKVEENTPDQITTPYIVVNGHIDPNSGDLALRSTLYG 180

DB 461 PENFEDVTSNLTFFPNQYKVEENTPDQITTPYIVVNGHIDPNSGDLALRSTLYG 520

QY 181 YNSNIIRWSMDSWNEVAFNNGSGSGDGIDKPVVPEQDE 219

DB 521 YNSNIIRWSMDSWNEVAFNNGSGSGDGIDKPVVPEQDE 559

XX RESULT 9

XX ABU16402

XX ID ABU16402 standard; protein; 935 AA.

XX XX

XX AC ABU16402;

QY 2 LTMPAYIDPENVKKTGNVTLATGIGSTANKTVLADYKEYKFNLSIKGTIDQIDKNN 61
Db 333 ITMPAYIDPENVTXGTNVTLTGIGTNTASKTVLIDYKEYGQFHNLSIKGTIDQIDKNN 392
QY 62 TYRQTIYVNPSSGDNVIAPIVLTGNLKPNVTDGNALIDQNTSIKVKYKVDNAADLSSEYFVNP 121
Db 393 TYRQTIYVNPSSGDNVIAPIVLTGNLKPNVTDGNALIDQNTSIKVKYKVDNAADLSSEYFVNP 452
QY 122 ENFEDVTNSVNIPTFPNPNQYKVEFTPDQITTTPIYVWNGHIDPNKGDALBSTLYGY 181
Db 453 SDFEDVTNQVRIISFPNANQYKVEFTPDQITTTPIYVWNGHIDPNKGDALBSTLYGY 512
QY 182 NSNIIWRSMWDNEVAFNNGSGDGDIDKPVVPEQDE 219
Db 513 DSNFIWRSMWDNEVAFNNGSGDGDIDKPVVPEQDE 550

RESULT 10
AAU33975
ID AAU33975 standard; protein; 1021 AA.
XX
AC AAU33975;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #251.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207272P.
PR 23-OCT-2000; 2000US-0243578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS51834.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 5471; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic

XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #1929.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA20272.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 44326; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC or a gene on which the test compound that inhibits proliferation of an
CC pathway in which a proliferation-required gene or its gene product lies
CC identifying a gene required for cellular proliferation or the biological
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 935 AA;
Query Match 85.7%; Score 991; DB 6; Length 935;
Best Local Similarity 84.9%; Pred. NO. 1.7e-72;
Matches 185; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1021 AA;

Query Match 85.7%; Score 991; DB 4; Length 1021;
 Best Local Similarity 84.9%; Pred. No. 1.9e-72;
 Matches 185; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTILVDYKGYKFNLSIKGTIDQDKTNN 61
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 333 ITMPAYIDPENVTGTGNVLTGIGTNTASKTILVDYKGYKFNLSIKGTIDQDKTNN 392

QY 62 TYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYYKVDNAADLSSEYFVNP 121
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 393 TYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYYKVDNAADLSSEYFVNP 452

QY 122 ENFEDVTNSVNIITFPNPNQYKVEFTDDQITTPYIVVNGHIDPNSKGDIALRSTLYGY 181
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 453 SDFEDVTNQVRISFPNPNQYKVEFTDDQITTPYIVVNGHIDPNSKGDIALRSTLYGY 512

QY 182 NSNIIRMSMSDNEVAFNNGSGSGDGIDKPVVPEQDPE 219
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 513 DSNFIWRMSDNEVAFNNGSGSGDGIDKPVVPEQDPE 550

RESULT 11
 ID AAU36951 standard; protein; 1021 AA.
 XX
 AC AAU36951;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus cellular proliferation protein #1121.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 PI Yamamoto RT, Xu RH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS54810.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 PT
 XX Example 3; SEQ ID NO 12544; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes

CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1021 AA;

Query Match 85.7%; Score 991; DB 4; Length 1021;
 Best Local Similarity 84.9%; Pred. No. 1.9e-72;
 Matches 185; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

QY 2 LTPAYIDPENVKKTGNVLTATGIGSTANKTILVDYKGYKFNLSIKGTIDQDKTNN 61
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 333 ITMPAYIDPENVTGTGNVLTGIGTNTASKTILVDYKGYKFNLSIKGTIDQDKTNN 392

QY 62 TYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYYKVDNAADLSSEYFVNP 121
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 393 TYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKYYKVDNAADLSSEYFVNP 452

QY 122 ENFEDVTNSVNIITFPNPNQYKVEFTDDQITTPYIVVNGHIDPNSKGDIALRSTLYGY 181
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 453 SDFEDVTNQVRISFPNPNQYKVEFTDDQITTPYIVVNGHIDPNSKGDIALRSTLYGY 512

QY 182 NSNIIRMSMSDNEVAFNNGSGSGDGIDKPVVPEQDPE 219
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 513 DSNFIWRMSMSDNEVAFNNGSGSGDGIDKPVVPEQDPE 550

RESULT 12
 ID AAU75490 standard; protein; 496 AA.
 XX
 AC AAU75490;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE S. aureus antigenic protein associated protein #10.
 XX
 KW Antigenic protein; vaccine; SEREX; antibacterial; antiinflammatory;
 KW dermatological; antiulcer; tuberculostatic; immunosuppressive;
 KW septicemia; food poisoning; skin disorders; peritonitis; endocarditis;
 KW tuberculosis; blood infection; sepsis; meningitis; pneumonia;
 KW stomach ulcer; gonorrhea; necrotising fasciitis; impetigo;
 KW Lyme's disease; gastro-enteritis; dysentery; shigellosis.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO200198499-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 20-JUN-2001; 2001WO-GB002685.
 XX
 PR 20-JUN-2000; 2000GB-00014907.
 XX
 PA (UYSH-) UNIV SHEFFIELD.
 PA (BIOS-) BIOSYNEXUS INC.
 XX
 PI Foster S, McDowell P, Brummell K, Clarke S;
 XX WPI; 2002-106544/14.
 DR
 XX

PT Identifying antigenic polypeptides expressed by pathogenic organisms
PT e.g., Staphylococcus aureus during infection, by SEREX (serological
PT identification of antigens by recombinant expression cloning) techniques.
XX
XX Disclosure; Page 72-74; 85pp; English.
PS
XX The invention relates to a method for identifying antigenic polypeptides
CC expressed by pathogenic organisms e.g., Staphylococcus aureus during
CC infection, by SEREX (serological identification of antigens by
CC recombinant expression cloning) techniques. The method involves providing
CC a nucleic acid library encoding genes/partial gene sequences of
CC pathogenic organisms, transforming/transfecting the library into host
CC cells, contacting the polypeptides expressed by the genes with autologous
CC antisera derived from an animal infected with, or has been infected with
CC the pathogen and purifying the nucleic acid encoding the polypeptide or
CC partial polypeptide binding to the antisera. Also included are the
CC nucleic acids and polypeptides isolated by the method, vectors and
CC transformed cells expressing them, a vaccine comprising the polypeptide
CC and the production of monoclonal antibodies against the polypeptides. The
CC protein and vaccine are useful for immunising an animal (preferably
CC human) against a pathogenic microbe. The proteins and antibodies are
CC useful for manufacturing a medicament for treating Staphylococcus aureus-
CC associated septicaemia, food poisoning or skin disorders or
CC Staphylococcus epidermidis-associated septicaemia, sepsis, meningitis,
CC endocarditis, tuberculosis, blood infections, sepsis, meningitis,
CC pneumonia, stomach ulcers, gonorrhea, necrotising fasciitis, impetigo,
CC Lyme's disease, gastro-enteritis, dysentery and shigellosis. The present
CC sequence is an S. aureus protein sequence associated with the antigenic
CC proteins of the invention. Note: The present sequence is included in the
CC sequence listing but is not mentioned anywhere else in the specification
XX
XX Sequence 496 AA;
SQ

Query Match 74.6%; Score 863; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.5e-62;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVLTATGISTTANKTVLVDYKYGKFNLSIKGTIDQDKTN 60
Db 332 TLTMPAYIDPENVKKTGNVLTATGISTTANKTVLVDYKYGKFNLSIKGTIDQDKTN 391
QY 61 NTVRQTIYVPSGDNVIAPIVLGNLKPNTDSNALIDQNTSIKVYKVDNAADLSESYFVN 120
Db 392 NTVRQTIYVPSGDNVIAPIVLGNLKPNTDSNALIDQNTSIKVYKVDNAADLSESYFVN 451
QY 121 PENFEDVTSNVIITFPNPQYKVEFNTPDQITTPYIVVNGHID 165
Db 452 PENFEDVTSNVIITFPNPQYKVEFNTPDQITTPYIVVNGHID 496

RESULT 13
ADA89664
ID ADA89664 standard; protein; 496 AA.
XX
AC ADA89664;
XX
DT 20-NOV-2003 (first entry)
XX
DE Staphylococcus aureus antigenic protein #203.
XX
XX antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;
KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;
KW bacteraemia; septic shock; organ infection; skin infection;
KW bacterial basal colonisation; bacterial eye infection; septicaemia;
KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;
KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
KW gastro-enteritis; dysentery; shigellosis; skin disorder.
XX
OS Staphylococcus aureus.
XX
XX WO2003011899-A2.
PN

XX
PD 13-FEB-2003.
XX
XX 02-AUG-2002; 2002WO-GB003606.
XX
PR 02-AUG-2001; 2001GB-00018825.
PR 09-JAN-2002; 2002GB-00000349.
XX (UYSH-) UNIV SHEFFIELD.
XX (BIOS-) BIOSYNEXUS INC.
XX Foster S, Mord J, Clarke S, Medowell P, Brummel K;
XX WPI; 2003-256434/25.
XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
XX useful as a vaccine for immunizing humans against e.g. bacteraemia, septic
XX shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhoea or
XX impetigo.
XX Claim 4; Page 161; 189pp; English.
XX The present invention describes an antigenic protein or its part, which
XX is for use as a vaccine. The antigenic protein is encoded by an isolated
XX DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
XX sequences (designated dnaSA and dna SE, respectively; and which encodes a
XX protein expressed by a pathogenic organism. Also described: (1) a vaccine
XX composition comprising at least one antigenic protein; (2) a method of
XX immunising an animal against a disease or condition caused by a
XX pathogenic microbe by administering the antigenic protein or the vaccine;
XX (3) an antibody or its binding part obtainable by the method above; (4)
XX preparing a hybridoma cell line producing monoclonal antibodies; (5) a
XX hybridoma cell line produced by the method of (4); and (6) identifying
XX opsonic antigens expressed by a pathogenic microbe. The antigenic
XX proteins have antibacterial, neuroprotective, immunosuppressive,
XX antiinflammatory, antiulcer, immunostimulant and ophthalmological
XX activities, and can be used in vaccines. The antigenic proteins or
XX vaccines can be used for immunising an animal (specifically a human)
XX against a disease or condition caused by a pathogenic microbe, e.g.
XX bacteraemia, septic shock, organ infection, skin infection, bacterial
XX basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
XX bacteria-associated food poisoning, blood infections, peritonitis,
XX endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
XX strep throat, streptococcal-associated toxic shock, necrotising
XX fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
XX dysentery, shigellosis, S. aureus-associated septicaemia, food-poisoning,
XX skin disorders, S. epidermidis-associated septicaemia, peritonitis or
XX endocarditis. The present sequence represents a S. aureus antigenic
XX protein sequence from the present invention.
XX
XX Sequence 496 AA;
SQ

Query Match 74.6%; Score 863; DB 6; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.5e-62;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLTMPAYIDPENVKKTGNVLTATGISTTANKTVLVDYKYGKFNLSIKGTIDQDKTN 60
Db 332 TLTMPAYIDPENVKKTGNVLTATGISTTANKTVLVDYKYGKFNLSIKGTIDQDKTN 391
QY 61 NTVRQTIYVPSGDNVIAPIVLGNLKPNTDSNALIDQNTSIKVYKVDNAADLSESYFVN 120
Db 392 NTVRQTIYVPSGDNVIAPIVLGNLKPNTDSNALIDQNTSIKVYKVDNAADLSESYFVN 451
QY 121 PENFEDVTSNVIITFPNPQYKVEFNTPDQITTPYIVVNGHID 165
Db 452 PENFEDVTSNVIITFPNPQYKVEFNTPDQITTPYIVVNGHID 496

RESULT 14
ADA89663
ID ADA89663 standard; protein; 496 AA.
XX
XX

AC ADAB9663;
 DT 20-NOV-2003 (first entry)
 XX
 DE Staphylococcus aureus antigenic protein #202.
 XX
 KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;
 KW antibacterial; neuroprotective; immunosuppressive; anti-inflammatory;
 KW antitumor; immunostimulant; ophthalmological; pathogenic microbe;
 KW bacteraemia; septic shock; organ infection; skin infection;
 KW bacterial basal colonisation; bacterial eye infection; septicaemia;
 KW tuberculous; food poisoning; blood infection; peritonitis; endocarditis;
 KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;
 KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;
 KW gastro-enteritis; dysentery; shigellosis; skin disorder.
 XX
 OS Staphylococcus aureus.
 XX
 XX WO2003011899-A2.
 PN
 XX 13-FEB-2003.
 XX
 XX 02-AUG-2002; 2002WO-GB003606.
 XX
 XX 02-AUG-2001; 2001GB-00018825.
 PR
 XX 09-JAN-2002; 2002GB-00000349.
 XX
 XX (UYSH-) UNIV SHEFFIELD.
 PA (BIOS-) BIOSYNEXUS INC.
 XX
 XX Foster S, Mond J, Clarke S, McDowell P, Brummel K;
 PI
 XX WPI; 2003-256434/25.
 DR
 XX
 XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,
 PT useful as a vaccine for immunizing humans against e.g. bacteremia, septic
 PT shock, septicaemia, tuberculosis, meningitis, pneumonia, gonorrhea or
 PT impetigo.
 XX
 XX Claim 4; Page 160-161; 189pp; English.
 PS
 XX The present invention describes an antigenic protein or its part, which
 CC is for use as a vaccine. The antigenic protein is encoded by an isolated
 CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene
 CC sequences (designated dnaA and dna SE, respectively; and which encodes a
 CC protein expressed by a pathogenic organism. Also described: (1) a vaccine
 CC composition comprising at least one antigenic protein; (2) a method of
 CC immunising an animal against a disease or condition caused by a
 CC pathogenic microbe by administering the antigenic protein or the vaccine;
 CC (3) an antibody or its binding part obtainable by the method above; (4)
 CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a
 CC hybridoma cell line produced by the method of (4); and (6) identifying
 CC opsonic antigens expressed by a pathogenic microbe. The antigenic
 CC proteins have antibacterial, neuroprotective, immunosuppressive,
 CC antiinflammatory, antitumor, immunostimulant and ophthalmological
 CC activities, and can be used in vaccines. The antigenic proteins or
 CC vaccines can be used for immunising an animal (specifically a human)
 CC against a disease or condition caused by a pathogenic microbe, e.g.
 CC bacteraemia, septic shock, organ infection, skin infection, bacterial
 CC basal colonisation, bacterial eye infections, septicaemia, tuberculosis,
 CC bacteria-associated food poisoning, blood infections, peritonitis,
 CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,
 CC strep throat, streptococcal-associated toxic shock, necrotising
 CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,
 CC skin disorders, S. aureus-associated septicaemia, food-poisoning,
 CC endocarditis. The present sequence represents a S. aureus antigenic
 CC protein sequence from the present invention.
 XX
 XX Sequence 496 AA;
 SQ
 Query Match 74.6%; Score 863; DB 6; Length 496;
 Best Local Similarity 100.0%; Pred. No. 2.5e-62;

Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 60
 DB |||||
 332 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTN 391
 QY 61 NTYRQTIYVNSGDNVIAPIVLTGNLKNPTSDNALIDQOQNTSIKVKVDNAADLSSSYFN 120
 DB |||||
 332 NTYRQTIYVNSGDNVIAPIVLTGNLKNPTSDNALIDQOQNTSIKVKVDNAADLSSSYFN 451
 QY 121 PENFEDVTNSNIITFPNPNQYKVPENTDDQITTPYIVVNGHID 165
 DB |||||
 452 PENFEDVTNSNIITFPNPNQYKVPENTDDQITTPYIVVNGHID 496

RESULT 15
 AAU34403
 ID AAU34403 standard; protein; 767 AA.
 XX
 AC AAU34403;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus cellular proliferation protein #679.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Staphylococcus aureus.
 XX
 XX WO200170955-A2.
 PN
 XX 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US009180.
 XX
 XX 21-MAR-2000; 2000US-0191078P.
 PR
 XX 23-MAY-2000; 2000US-0206848P.
 PR
 XX 26-MAY-2000; 2000US-0207727P.
 PR
 XX 23-OCT-2000; 2000US-0242578P.
 PR
 XX 27-NOV-2000; 2000US-0253625P.
 PR
 XX 22-DEC-2000; 2000US-0257911P.
 PR
 XX 16-FEB-2001; 2001US-0269308P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 RI Yamamoto RT, Xu HH;
 XX
 XX WPI; 2001-611495/70.
 DR
 XX N-PSNB; AAS52262.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 XX Example 3; SEQ ID NO 5899; 51pp; English.
 XX
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 767 AA;

```
Query Match          22.0%; Score 255; DB 4; Length 767;
Best Local Similarity 26.9%; Pred. No. 3.6e-12;
Matches 61; Conservative 54; Mismatches 82; Indels 30; Gaps 8;

QY      2 LTMFAYIDPENVKKTGNVTLATGIGSTTANKTVLVY-----EYXGKPYNLSIKGTIDQI 56
Db      273 LSLNLFIDPTVTQKGNQVEVKLGERTVSKIFNIQYLGVRDNGW----VTANGRIDTL 328
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      57 DKTNTYRQTLTYVNPSSGDNVIAPVLTGNL-----KPNTDSNALIDQONTISIKVYKVDNAAD 112
Db      329 NKVDGKTSHPAYMKPNQSLSSVTGTQVTKGNKPGVN-----NPTVKYKHIGSDDD 380
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      113 LSESYFVNPN-----FEDVTSNVNITFPNPNOYKVEFNTPDDQITTPYIVVVVNGHIDPNSK 169
Db      381 LAESVYAKLDDVSKFEDVTDNWSLDFDTNGGYSLNFNLDQ--SKYVYIKYEGYDSNA- 437
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      170 GDLALRSTLYGYNINIIWRMSWDNEVAFNNGSGSGDGIDK---PVV 213
Db      438 SNLBFQTHLFGYNYVYVYTSNLTWKNGVAFYSNNAQGGKDKLKEPII 484
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Search completed: November 3, 2004, 19:03:02
Job time : 41.4791 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 18:54:26 ; Search time 10.7401 Seconds
(without alignments)
1352.277 Million cell updates/sec

Title: US-09-679-643-2_COPY_332_550

Perfect score: 1157

Sequence: 1 TLTMPAYIDPENVKKTGNVT.....NGSGSGDGIDKPVVPEQFDE 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1157	100.0	345	US-08-856-253-7	Sequence 7, Appli
2	1157	100.0	933	US-08-293-728-2	Sequence 2, Appli
3	1157	100.0	933	US-09-421-868-2	Sequence 2, Appli
4	1157	100.0	936	US-08-956-171E-5249	Sequence 5249, Ap
5	1157	100.0	936	US-08-781-986A-5249	Sequence 5249, Ap
6	244.5	21.1	930	US-09-134-001C-5314	Sequence 5314, Ap
7	244.5	21.1	930	US-09-386-962C-10	Sequence 10, Appl
8	241.5	20.9	582	US-09-147-405B-13	Sequence 13, Appl
9	241.5	20.9	593	US-09-147-405B-11	Sequence 11, Appl
10	241.5	20.9	1092	US-09-147-405B-15	Sequence 15, Appl
11	208	18.0	918	US-09-200-650E-7	Sequence 7, Appli
12	207.5	17.9	1166	US-09-200-650E-11	Sequence 11, Appl
13	180.5	15.6	1315	US-09-200-650E-5	Sequence 5, Appli
14	175.5	15.2	1027	US-08-956-171E-5254	Sequence 5254, Ap
15	175.5	15.2	1027	US-08-781-986A-5254	Sequence 5254, Ap
16	136	11.8	559	US-08-956-171E-5251	Sequence 5251, Ap
17	136	11.8	559	US-08-781-986A-5251	Sequence 5251, Ap
18	128.5	11.1	251	US-08-956-171E-5252	Sequence 5252, Ap
19	128.5	11.1	251	US-08-781-986A-5252	Sequence 5252, Ap
20	123	10.6	1112	US-08-714-402-2	Sequence 2, Appli
21	123	10.6	1161	US-09-327-536-2	Sequence 2, Appli
22	117.5	10.2	930	US-09-200-650E-3	Sequence 3, Appli
23	112	9.7	1155	US-09-710-279-1780	Sequence 1780, Ap
24	112	9.7	1742	US-09-386-962C-4	Sequence 4, Appli
25	112	9.7	1742	US-09-386-959-4	Sequence 4, Appli
26	104.5	9.0	1216	US-09-134-000C-5130	Sequence 5130, Ap
27	104.5	9.0	2780	US-10-220-587-2	Sequence 2, Appli

28	104	9.0	1132	4	US-09-248-796A-15026	Sequence 15026, A
29	103	8.9	405	4	US-09-328-352-5300	Sequence 5300, Ap
30	102.5	8.9	559	4	US-09-710-279-1376	Sequence 1376, Ap
31	100.5	8.7	1454	4	US-09-673-896-2	Sequence 2, Appli
32	98	8.5	1457	4	US-09-673-896-4	Sequence 4, Appli
33	97.5	8.4	186	2	US-08-732-791-1	Sequence 1, Appli
34	97.5	8.4	186	3	US-09-205-049-1	Sequence 1, Appli
35	97.5	8.4	1180	4	US-09-206-942-65	Sequence 65, Appl
36	97.5	8.4	1188	4	US-09-206-942-63	Sequence 63, Appl
37	94.5	8.2	992	4	US-09-206-942-61	Sequence 61, Appl
38	94.5	8.2	998	4	US-09-206-942-59	Sequence 59, Appl
39	94.5	8.2	1359	4	US-09-134-000C-6374	Sequence 6374, Ap
40	94	8.1	1095	4	US-09-206-942-45	Sequence 45, Appl
41	94	8.1	1101	4	US-09-206-942-43	Sequence 43, Appl
42	93.5	8.1	526	2	US-08-853-659A-40	Sequence 40, Appl
43	93.5	8.1	2777	4	US-10-220-587-4	Sequence 4, Appli
44	93	8.0	668	4	US-09-277-431A-2	Sequence 2, Appli
45	92.5	8.0	1228	4	US-09-463-402-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

Query Match 100.0%; Score 1157; DB 3; Length 345;
Best Local Similarity 100.0%; Pred.No. 3.4e-105;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. NO. 1.5e-104; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 125 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 184

QY 61 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 120
DB 185 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 244

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 180
DB 245 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 304

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 305 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 343

RESULT 2
US-08-293-728-2
Sequence 2, Application US/08293728D
Patent No. 6008341
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/08/293,728D
CURRENT FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. NO. 1.5e-104; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 391

QY 61 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 120
DB 392 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 451

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 180
DB 452 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 511

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 512 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 550

RESULT 4
US-08-956-171E-5249
Sequence 5249, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 125 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 184

QY 61 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 120
DB 185 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 244

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 180
DB 245 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 304

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 305 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 343

RESULT 2
US-08-293-728-2
Sequence 2, Application US/08293728D
Patent No. 6008341
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/08/293,728D
CURRENT FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 100.0%; Score 1157; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. NO. 1.5e-104; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0;

QY 1 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 60
DB 332 TLTPAYIDPENVKKTGNVLTATGIGSTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 391

QY 61 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 120
DB 392 NTRYQTIYVNPSPGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYVDNAADLSSEYFVN 451

QY 121 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 180
DB 452 PENFEDVTNSVNIITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNSKGDALALRSTLYG 511

QY 181 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 219
DB 512 YNSNIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQDE 550

RESULT 3
US-09-421-868-2
Sequence 2, Application US/09421868
Patent No. 6177084
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/09/421,868
CURRENT FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 08/293,728
PRIOR FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5249:
US-08-956-171E-5249

Query Match          100.0%; Score 1157; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDIDKTN 60
Db 341 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDIDKTN 400

QY 61 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFN 120
Db 401 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFN 460

QY 121 PENFEDVTSNVIITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 180
Db 461 PENFEDVTSNVIITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 520

QY 181 YNSNIIWRSMWDEVAFNNGSGDGDIDKPVVPEQDPE 219
Db 521 YNSNIIWRSMWDEVAFNNGSGDGDIDKPVVPEQDPE 559

RESULT 5
US-08-781-986A-5249
; Sequence 5249, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB2499P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5249:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 936 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match          100.0%; Score 1157; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDIDKTN 60
Db 341 TLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDIDKTN 400

QY 61 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFN 120
Db 401 NTRYQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYFN 460

QY 121 PENFEDVTSNVIITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 180
Db 461 PENFEDVTSNVIITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGLALRSTLYG 520

QY 181 YNSNIIWRSMWDEVAFNNGSGDGDIDKPVVPEQDPE 219
Db 521 YNSNIIWRSMWDEVAFNNGSGDGDIDKPVVPEQDPE 559

RESULT 6
US-09-134-001C-5314
; Sequence 5314, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5314
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match          21.1%; Score 244.5; DB 3; Length 930;
Best Local Similarity 31.9%; Pred. No. 4.3e-15;
Matches 72; Conservative 36; Mismatches 93; Indels 25; Gaps 10;

QY 2 LTWPAYIDPENVKKTG---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDIDK 58
Db 386 LKTSYIDKSKVNNNTKLDVEYKTALSS--VNKTITVEYQKPNENRTANLQSMFTNIDT 443

QY 59 TNNYRQTIYVNPSSGDNVIAPIVLTGNLKPNTDSNALIDQNTSIKVKVDNAADLSESYF 118
Db 444 KKHTEQTIYINPL--RYSAKETNVIISGNDGEGSTIIDSTIIKVKVGNQNLFPDSNR 501

QY 119 V-NPENFEDVTSNVIITFPNPQYKVFNTPDQDITTPYIVVNGHIDPNSKGD----- 171
Db 502 IYDYSEYEDVTNDDYAQLGNNDVNINFG---NIDSPVIKVISKYDPN-KDYTTIIQ 556

QY 172 -LALRSTLYGNSNIIWRSMWDEVAFNNGSGDGDIDKPVVPEQ 216
Db 557 TVTMQTTINEYTG--FRTASYDNTIAFTSTSSGQGG--DLP--PEK 597

RESULT 7
US-09-386-962C-10
; Sequence 10, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; CURRENT FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
```

```

; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-386-962C-10

Query Match      21.1%; Score 244.5; DB 4; Length 930;
Best Local Similarity 31.9%; Pred. No. 4.3e-15;
Matches 72; Conservative 36; Mismatches 93; Indels 25; Gaps 10;

Qy 2 LTPPAYIDPENVKKTG---NVTLATGIGSTTANKTVLDVDEYKGFYKFNLSIKGTIDQIDK 58
Db 386 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRNTANLQSMFTNIDT 443
Qy 59 TNNTYRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 444 KNHIVEQTIYINPL--RYSAKETNVNISNGDEGSTIIDDSITIIKVKYKVDNQNLPSNR 501
Qy 119 V-NPENFEDVTNSVNIITPPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGD----- 171
Db 502 IYDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVIKSKYDPN-KDDYTTIIQQ 556
Qy 172 -LALRSTLYGYSNIIWRSMWSDNEVAFNNGSGDGDIDKPVVPEQ 216
Db 557 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG--DLP--PEK 597

```

```

RESULT 8
US-09-147-405B-13
; Sequence 13, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09/147,405B
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-13

Query Match      20.9%; Score 241.5; DB 4; Length 582;
Best Local Similarity 31.4%; Pred. No. 4.2e-15;
Matches 71; Conservative 37; Mismatches 93; Indels 25; Gaps 10;

Qy 2 LTPPAYIDPENVKKTG---NVTLATGIGSTTANKTVLDVDEYKGFYKFNLSIKGTIDQIDK 58
Db 315 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRNTANLQSMFTNIDT 372
Qy 59 TNNTYRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 373 KNHIVEQTIYINPL--RYSAKETNVNISNGDEGSTIIDDSITIIKVKYKVDNQNLPSNR 430
Qy 119 V-NPENFEDVTNSVNIITPPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGD----- 171
Db 431 IYDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVIKSKYDPN-KDDYTTIIQQ 485

```

```

Qy 172 -LALRSTLYGYSNIIWRSMWSDNEVAFNNGSGDGDIDKPVVPEQ 216
Db 486 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG--DLP--PEK 526

RESULT 9
US-09-147-405B-11
; Sequence 11, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09/147,405B
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-147-405B-11

Query Match      20.9%; Score 241.5; DB 4; Length 593;
Best Local Similarity 31.4%; Pred. No. 4.3e-15;
Matches 71; Conservative 37; Mismatches 93; Indels 25; Gaps 10;

Qy 2 LTPPAYIDPENVKKTG---NVTLATGIGSTTANKTVLDVDEYKGFYKFNLSIKGTIDQIDK 58
Db 322 LKLTSYIDKSKVPNNNTKLDVEYKTALSS--VNKTITVEYQRPENRNTANLQSMFTNIDT 379
Qy 59 TNNTYRQTIYVNPBGDNVIAPLVTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSSEYF 118
Db 380 KNHIVEQTIYINPL--RYSAKETNVNISNGDEGSTIIDDSITIIKVKYKVDNQNLPSNR 437
Qy 119 V-NPENFEDVTNSVNIITPPNPQYKVEFTPDQITTPYIVVNGHIDPNSKGD----- 171
Db 438 IYDYSEYEDVTNDYDQAQGNNDVNINFG---NIDSPYIIKVIKSKYDPN-KDDYTTIIQQ 492
Qy 172 -LALRSTLYGYSNIIWRSMWSDNEVAFNNGSGDGDIDKPVVPEQ 216
Db 493 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG--DLP--PEK 533

```

```

RESULT 10
US-09-147-405B-15
; Sequence 15, Application US/09147405B
; Patent No. 6733758
; GENERAL INFORMATION:
; APPLICANT: Guss, Bengt
; APPLICANT: Nilsson, Martin
; APPLICANT: Frykberg, Lars
; APPLICANT: Flock, Jan-Ingmar
; APPLICANT: Lindberg, Martin
; TITLE OF INVENTION: Fibrinogen Binding Protein Originating from
; FILE REFERENCE: guss 09/147405
; CURRENT APPLICATION NUMBER: US/09/147,405B
; PRIOR FILING DATE: 1999-04-11
; PRIOR APPLICATION NUMBER: PCT/SE97/10191
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: SE 9602496-3
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 15

```


This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1157	100.0	933	2	S41539	fibronogen-binding	
2	991	85.7	989	2	D89852	fibronogen-binding	
3	255	22.0	940	2	S19702	fibronectin-binding	
4	241.5	20.9	1092	2	T30214	fibronogen-binding	
5	237.5	20.5	961	2	G90053	hypothetical prote	
6	208	18.0	877	2	F90070	Clumping factor B	
7	207.5	17.9	1166	2	T28680	fibronogen-binding	
8	200.5	17.3	1141	2	E98824	hypothetical prote	
9	180.5	15.6	1315	2	T28679	fibronogen-binding	
10	175.5	15.2	1018	2	A32192	fibronectin-binding	
11	160.5	13.9	1038	2	H90053	hypothetical prote	
12	160.5	13.9	1385	2	D89824	hypothetical prote	
13	122.5	10.6	953	2	C89824	hypothetical prote	
14	110	9.5	1039	2	T30856	protein F2 - Strept	
15	109.5	9.5	1386	2	AC1533	surface protein (L	
16	103	8.9	421	2	AC1541	hypothetical prote	
17	102.5	8.9	3283	2	AC1018	large repetitive p	
18	102	8.8	294	2	T05604	hypothetical prote	
19	102	8.8	5005	2	F82884	hypothetical prote	
20	101.5	8.8	655	2	A57681	hypothetical prote	
21	101	8.7	642	2	D81401	probable flagellar	
22	101	8.7	4152	2	T31102	filamentous hemagg	
23	99	8.6	439	2	AE1251	probable peptidogl	
24	98	8.5	691	2	B75622	hypothetical prote	
25	98	8.5	1457	2	D81019	adhesion and penet	
26	98	8.5	2269	2	T18472	hypothetical prote	
27	97.5	8.4	463	2	A44808	cellulase (EC 3.2.	
28	97	8.4	676	2	C37775	acylamino-acid-rel	
29	97	8.4	749	2	S64788	hypothetical prote	

A:Molecule type: DNA
A:Residues: 1-989 <KUR>
A:Cross-references: UNIPROT:Q99VJ4; GB:BA000018; PID:gl3700678; PIDN:BAB41975.1; GSPDB:C
A:Experimental source: strain N315
C:Genetics:
A:Gene: Clfa

Query Match 85.7%; Score 991; DB 2; Length 989;
Best Local Similarity 84.9%; Pred. No. 6.6e-65;
Matches 185; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

Qy 2 LTMFAIDPENVKKTGNVLTATGISTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 61
Db 333 ITMFAIDPENVKKTGNVLTATGISTTANKTVLVDYKGYKFNLSIKGTIDQIDKTN 392

Qy 62 TYRQTIYVNPVPSGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLSESYFNP 121
Db 393 TYRQTIYVNPVPSGDNVIAPIVLTGNLKPNTKSNALIDAKNTDIKVKYVDNANDLSESYFNP 452

Qy 122 ENFEDVTVNSVMTFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSKGLALRSTLYGY 181
Db 453 SDFEDVTVNQVRSFPNAYKVEFNTPDQITTPYIVVNGHIDPASTGDALRSTTFGY 512

Qy 182 NSNIWRSMWNEVAFNNGSGDGIDKPVVPEQDE 219
Db 513 DSNFTWRSMWNEVAFNNGSGDGIDKPVVPEQDE 550

RESULT 3
S19702
fibronectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004
R:Accession: S19702
R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus
A:Reference number: S19702; MUID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <JOE>
A:Cross-references: UNIPROT:Q53682; EMBL:X62992; NID:G49040; PIDN:CAA44726.1; PID:G58156
C:Keywords: fibronectin binding

Query Match 22.0%; Score 255; DB 2; Length 940;
Best Local Similarity 26.9%; Pred. No. 7.1e-11;
Matches 61; Conservative 54; Mismatches 82; Indels 30; Gaps 8;

Qy 2 LTMFAIDPENVKKTGNVLTATGISTTANKTVLVDY----EKYKGFNLSIKGTIDQI 56
Db 273 LSNLFIDPTTVTKGNQNVKLGTTVSKLFNIOYLGVRDNG----VTANGRIDTL 328

Qy 57 DKTNNYTRQTIYVNPVPSGDNVIAPIVLTGNL----KPNTDSNALIDQOQNTSIKVKYKVDNAAD 112
Db 329 NKVDGKSFHFAYMKPNQSLSSVTVTGQYTKGNKPGV-----NPTVKVYKHIGSD 380

Qy 113 LSESYFNPEN---FEDVTVNSVMTFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSK 169
Db 381 LAESVYAKLDVSKFEDVTVNSMELDFDTNGGYSLENFNLDQ---SKNYVIKGYGYDSNA- 437

Qy 170 GDALALRSTLYGNSNIWRSMWNEVAFNNGSGDGIDK---PVV 213
Db 438 SNLEFQTHLFGYYNYTSLTWNGVAFYSNNAQDGDKLEPII 484

RESULT 4
T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C:Date: 22-Oct-1999 #sequence revision 22-Oct-1999 #text_change 09-Jul-2004
A:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: Z20781; MUID:98261511; PMID:9596732
A:Accession: T30214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <NIL>
A:Cross-references: UNIPROT:O70022; EMBL:Y17116; NID:el296734; PID:el296735; PIDN:CAA766

Query Match 20.9%; Score 241.5; DB 2; Length 1092;
Best Local Similarity 31.4%; Pred. No. 8.5e-10;
Matches 71; Conservative 37; Mismatches 93; Indels 25; Gaps 10;

Qy 2 LTMFAIDPENVKKTG---NVLTATGISTTANKTVLVDYKGYKFNLSIKGTIDQIDK 58
Db 389 LKLTYSIDKSKVNNNTKLDVEYKLTALSS--VNKTIITVEYQRPNERNTANLQSMFTNIDT 446

Qy 59 TNNYTRQTIYVNPVPSGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLSESYF 118
Db 447 KNHTVEQTIYINPL--RYSAKETNNVNSGNGDEGTSIIDDSTIIIKVYKVGONQNLFDNSR 504

Qy 119 V-NPENPEFVTVNSVMTFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSKGD----- 171
Db 505 IYDSEYEDVTNDYQAQGNNDVYNIFG---NIDSPYIIKVISKYDEN-KDDVTTIQQ 559

Qy 172 -LALRSTLYGNSNIWRSMWNEVAFNNGSGDGIDKPVVPEQ 216
Db 560 TVTMQTITNEYTGE--FRTASYDNTIAFSTSSGQGG-DLP--PEK 600

RESULT 5
G90053
hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
A:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A9758; MUID:21311952; PMID:11418146
A:Accession: G90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-961 <KUR>
A:Cross-references: UNIPROT:Q99RD3; GB:BA000018; PID:gl3702452; PIDN:BAB43593.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnbB

Query Match 20.5%; Score 237.5; DB 2; Length 961;
Best Local Similarity 26.6%; Pred. No. 1.4e-09;
Matches 61; Conservative 47; Mismatches 94; Indels 27; Gaps 7;

Qy 2 LTMFAIDPENVKKTGNVLTATGISTTANKTVLVDY-----EKYKFNLSIKGTIDQI 56
Db 276 LNLNLFIDPTTVTKQGSQKVEVTLGONKYSKEFDIKYLDGVKDRMG----VTYNGRIDTL 331

Qy 57 DKTNNYTRQTIYVNPVPSGDNVIAPIVLTGNLKPNTDSNALIDQOQNTSIKVKYKVDNAADLSE 116
Db 332 NKEEGKSFHFAYKVPKNQSLTSVTVTGQVTSYGKQSA---NNPTVKVYKHIGSDLAES 387

Qy 117 YFV---NPNFEDVTVNSVMTFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSKGLA 173
Db 388 VYAKLDVTSKFEVDVTEKVNLSYTSNGGYTLNLGLDNL--SKDYVIKGYEYDQNAK-DLN 444

Qy 174 LRSTLYGNSNIWR-----SMWNEVAFNNGSGDGIDKPVV 214
Db 445 FRTHLSGYHKYFYPYYPYYPYVQVLTWNGVAFYSNNAKGDKKNDP 493

RESULT 6

F90070
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90070
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <KUR>
A:Cross-references: UNIPROT:Q99R07; GB:BA000018; PID:g13702588; PIDN:BA043728.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfB

Query Match 18.0%; Score 208; DB 2; Length 877;
Best Local Similarity 25.8%; Pred. No. 1.8e-07;
Matches 60; Conservative 46; Mismatches 99; Indels 28; Gaps 10;

QY 3 TMAPYIDPENVKKTGNVTLATGISTTANKTVLVDEK-----YCKPYNLSIKGTIDQDK 58
Db 324 SLEFLTDRAKAPKSGTYDANINIADEMFNNKITYNSSPIAGIDKPGANISSQIIIGVDT 383
QY 59 TN--NTYRQTIYVNP-----GDNVIAPLVLTGNLKPNTDSNALIDQNTSIKVYKVDNAAD 112
Db 384 ASQNTYIKQTVFVNPQRVLGNTWV--YIKGYQDKIEESGKVSATDTKLRIFEVNDTSK 441
QY 113 LSESYFVNP--ENFEDVTSV--NITFPNPQYKVFNTPDQITTPYIVVNGHIDPNS 168
Db 442 LSDSYVADPNDSLKEVTDQFKNRIYVEHNVASIKFG---DITKYVVLVEGHVDNTG 497
QY 169 KGLALRSTLYGN-----SNIIWRSMWDNEVAFNNGSGSGGIDKPVVPEQP 217
Db 498 KN---LKTQIQENVDFTNRDYSIFGNWENVRYGGGSADG--DSAVNPKDP 546

RESULT 7
T28680
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28680
R:Joseffson, E.; McCrea, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of St
A:Reference number: T20510; MUID:99098700; PMID:9884231
A:Accession: T28680
A>Status: preliminary; translated from GB/EMBL/DDBL
A:Molecule type: DNA
A:Residues: 1-1166 <JOS>
A:Cross-references: UNIPROT:O86489; EMBL:AJ005647; NID:e1318793; PID:e1318794; PIDN:CAAG
C:Genetics:
A:Gene: sdre

Query Match 17.9%; Score 207.5; DB 2; Length 1166;
Best Local Similarity 31.7%; Pred. No. 2.9e-07;
Matches 73; Conservative 36; Mismatches 88; Indels 33; Gaps 12;

QY 2 LTMPAYIDPNV--KKTGNVTLATGISTTANKTVLVDEYKGYKFNLSIKGTIDQDKT 59
Db 387 LTIYSYIDKKTVEVNTSLNLTFAAGKESQN--VTVDYQDPMVHGDSNIQSIFTKLDED 444
QY 60 NNTYRQTIYVNP-----SGDNVIAPLV-----TGNLKPNTDSNALIDQNTSIKVYKVDNA 110
Db 445 KQIEQIQIYVNPPLKKSATNKTVDIAGSQVDDYGNIKLNGS--TIID-QNTEIKYKVNDS 502
QY 111 ADLSESYFV--NPENFEDVTSVNTITFPNPQYKVFNTPD--DQITTPYIVVNGHIDPNS 168

Db 503 QQLPQSNRIYDFSQVEDVTQSQ-----FDNKKKSFENNVAITLDFGDINSAYIIKVVSXYTPTS 558

Qy 169 KGDL-----ALRST-LVGYNSNIWRSMWDNEAFNNNGSGSDGIDKP 211
|||:::|||||::| : | : | : |||
Db 559 DGEILDIAOGTSMTTKDYGX-----YNYAGYSNFIVTSDNTGGDGTVXP 603

RESULT 8

E89824
hypothetical protein sdrE [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: E89824
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi,
ma, A.; Mizutani-Gi, Y.; Kobayashi, N.; Sawano, R.; Inoue, R.; Kaito, C.; Sekimizu, K.; I
C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; PMID:21311952; PMID:11418146
A/Accession: E89824
A:Molecule type: DNA
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1141 <KUR>
A/Cross-references: UNIPROT.Q99W46; GB.BA000018; PID:g13700455; PIDN:BAB41752.1; GSPDB.G:
A/Experimental source: strain N315
C/Genetics:
A/Gene: sdrE

Query Match 17.3%; Score 200.5; DB 2; Length 1141;
Best Local Similarity 31.3%; Pred.No. 9.3e-07;
Matches 72; Conservative 36; Mismatches 89; Indels 33; Gaps 12;

Qy 2 LTMPAYIDPENV--KKTGNVTLATIGGSTTANKTVLDVEKYGFYNLSIKGTIDQIDKT 59
|||:::|||||::| : | : | : |||
Db 382 LTLXSYYDKQAVPNETSNLTFATAGKETSON--VSVDYQDPMVHGDSNIQIFTKLDEN 439

Qy 60 NNTRYQTIFYNP---SGDNVIAPVL-----TGNIKPMNTSNAIIDQQNTSIKYVKVDNA 110
|||:::|||||::| : | : | : |||
Db 440 QTIEQQIIVYVPLKKTATNTKVIDAGSQVDYGNIKLGNGS-TIID-QNTEIKVYKVNPN 497

Qy 111 ADLSESYPV-NPEPFEDVTNSVINTFEPNQYKFVFETPD-DQITTPYIVVVGHDIPS 168
|||:::|||||::| : | : | : |||
Db 498 QQLPQSNRIYDFSQVEDVTQSQ-----FDNKKSFENNVAITLDFGDINSAYIIKVVSXYTPTS 553

Qy 169 KGDL-----ALRST-LVGYNSNIWRSMWDNEAFNNNGSGSDGIDKP 211
|||:::|||||::| : | : | : |||
Db 554 DGEILDIAOGTSMTTKDYGX-----YNYAGYSNFIVTSDNTGGDGTVXP 598

RESULT 9

T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C/Species: Staphylococcus aureus
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28679
R/Josefsson, E.; McCrear, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.S
Microbiology 144, 3387-3395, 1998
A>Title: Three new members of the serine-aspartate repeat protein multigene family of Sta
A/Reference number: Z20510; PMID:99098700; PMID:9884231
A/Accession: T28679
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1315 <JOS>
A/Cross-references: UNIPROT.O84688; EMBL.AJ005646; NID:e1318791; PID:e1318792; PIDN:CRAO0
C/Genetics:
A/Gene: sdrD

Query Match 15.6%; Score 180.5; DB 2; Length 1315;
Best Local Similarity 27.7%; Pred.No. 3.3e-05;
Matches 58; Conservative 35; Mismatches 100; Indels 19; Gaps 8;

Qy 7 YIDPENVKTCGN-VLTLAGTIGSTTANKTVLDVEKYGFYNLSIKG----TIQDIQKTNN 61
|||:::|||||::| : | : | : |||

```
Db 361 YMDADTIPVSKNDVEFNVTIGNTTTKTIANIQPYDVVNEKNSIGSAFTETVSHVGNKEN 420
QY 62 T--YRQTIYVNPNGDNVIAPIVLTGNLKENTDSNAL--IDQNTSISKVYKDVAADLSBSY 117
Db 421 PYYQKTIYVNPSENLSNAKLKVOAYHSSYPNNIGQINKDVTDIKIYQPKGYTLNKG 480
QY 118 FVNPENFEDVNSV--NITFPNPQYKVEFNTPDQITTPYIVVNGHID--PNSKGDAL 174
Db 481 DVNTEKLTVDVNOYLQKITYGDNNSAVIDFGNAD----SAYVVMWNTKPYQNSPSPIV 536
QY 175 RSTLYGYSNIIWRSMWDNEVAFNNGSGSDG 207
Db 537 QMATLSSTGN--XSVSTGNALGFTNNOGGAG 566

RESULT 10
A32192
hypotheical protein - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997
C:Accession: A32192
R:Signaes, C.; Kaucii, G.; Joensson, K.; Lindgren, P.E.; Anantharamaiah, G.M.; Hoesek, M
Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989
A:Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylo
A:Reference number: A32192; MUID:89098998; PMID:2521391
A:Accession: A32192
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1018 <SIG>
A:Cross-references: GB:J04151
C:Keywords: fibronectin binding

Query Match 15.2%; Score 175.5; DB 2; Length 1018;
Best Local Similarity 19.6%; Pred. No. 5.5e-05; Mismatches 97; Indels 23; Gaps 9;
Matches 44; Conservative 61;

QY 2 LTMPAYIDPENVKKTGNVLTAGIGSTTANKTVLVVDY-EYKGFYNLSIKGTIDQIKTN 60
Db 301 LEINFIDPKTVQTNQNTITSTLNEEQTSKELDKYKDGIGNY-ANLNGSIETFNKAN 359
QY 61 NYRQTIYVNPNGDNVIAPIVLTGNLKENTDSNALIDQNTSISKVYK-VDNAADLSBSYFV 119
Db 360 NRFSHVAFKPNNGKTTSTVTGTLMKGSNQNG---NQPKVRIPEYLGNNEDIAKSVYA 415
QY 120 N---PENEDVTNSV--NITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSKGDAL 174
Db 416 NTDTSKFKFVTSNMSGNLNQNGSYSUNI----ENLDKTYVHVHDGEY-LNGTDEVDF 470
QY 175 RSTLYGYSNII-----WRSMWDNEVAFNNGSGSDGIDKPVV 213
Db 471 RTQMGHPQLKYKYDYRGYTLTDWNLGLYSNKANGNEKNGPII 515

RESULT 11
H90053
hypotheical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A99758; MUID:21311952; PMID:11418146
A:Accession: H90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: UNIPROT:Q99RD2; GB:BA000018; PID:g13702453; PIDN:BAB43594.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb
```

```
Query Match 13.9%; Score 160.5; DB 2; Length 1038;
Best Local Similarity 22.8%; Pred. No. 0.00071;
Matches 53; Conservative 49; Mismatches 107; Indels 23; Gaps 10;

QY 2 LTMPAYIDPENVKKTGNVLTAGIGSTTANKTVLVVDYKGYKFNLSIKGTIDQIKTN 61
Db 295 LEINFIDPKTVQSNGEQKITSKLNGEETKTIPIVYVNPNGVSNSTYVNGSIETENKSN 354
QY 62 TYRQTIYVNPNGDNVIAPIVLTGNLKENTDSNALIDQNTSISKVYK-VDNAADLSBSYFV 119
Db 355 KFTIAYIKPMNGQSNSTVSTGLT---TEGSLAGQPT-VKVYFYLGKKDELQSVYA 410
QY 120 NPEN---PEDVTNSVN--ITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNSKGDAL 174
Db 411 NTSOTNKEFVKEMNGKLSVQDNGSYSINL----DKLDTYVHVHTGEYLQGS-DQVNF 465
QY 175 RSTLYG-----YNSNIIWRSMWDNEVAFNNGSGSDGIDKPVVPPQDFE 219
Db 466 RTELGYPERAYKSYVYGGYRLTWDNLGLVYSNKAADGNGKNGQIIQDNDFE 517
```

```
RESULT 12
D99824
hypotheical protein sdrB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D99824
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A99758; MUID:21311952; PMID:11418146
A:Accession: D99824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1385 <KUR>
A:Cross-references: UNIPROT:Q99W47; GB:BA000018; PID:g13700454; PIDN:BAB41751.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: sdrB
```

```
Query Match 13.9%; Score 160.5; DB 2; Length 1385;
Best Local Similarity 25.1%; Pred. No. 0.001;
Matches 60; Conservative 37; Mismatches 99; Indels 43; Gaps 10;

QY 1 LTMPAYIDPENVKKTG-----NVLTATGIGSTTANKTVLVVDYE 39
Db 339 TYTFTDYVDRFNSVRKGINYSIYMDADTIPVDKDVFPFSTIGNQITTTAD----ITYP 394
QY 40 KYGKFYNLSIKG---TIDQIDKTNNT--YRQTIYVNPNGDNVI-APVLTGNLKENTDSN 92
Db 395 AYKEADNNSIGSAFTETVSHVGNVEDPGYVNVNPMKDLKGAKLKVAYHPKYPIN 454
QY 93 -ALIDQNTSISKVYKDVAADLSBSYFVNPENFEDVNSV--NITFPNPQYKVEFNTPD 149
Db 455 IQLINQNVNIIKIVRPEGYTLNKGVDVNTDLVDVTEBFKNKMTYGSNQSVNLDGFG--- 511
QY 150 DQITTPYIVVNGHID-PNSKGDALRSTLYGYSNIIWRSMWDNEVAFNNGSGSDG 207
Db 512 -DITSAYVVMWNTKPYQNSPSPIVQMATLSSTGN---KSVSTGNALGFTNNOGGAG 566
```

```
RESULT 13
C89824
hypotheical protein sdrC [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C89824
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
```


Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: C89824
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-953 <RUR>
A;Cross-references: UNIPROT:Q99W48; GB:BA000018; PID:gl3700453; PIDN:BA81750.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: sdrC

Query Match 10.6%; Score 122.5; DB 2; Length 953;
Best Local Similarity 23.4%; Pred. No. 0.39;
Matches 54; Conservative 32; Mismatches 108; Indels 37; Gaps 8;

Qy	1	TLTWPAYIDP-ENVK-----KTGNVT-----LATGIGSTTANKVLVDYKYGK 43
Db	281	TYTPTNVDDQYTNVSGSFEQVAFAKRENATTDKTAYRMEVTLGNDTYSKDVIVDYGKQG 340
Qy	44	FYNLSIKGTIDQIDKTNNTYRQTLTYVNPSGDNVIAVLGNL---KPNTDSNALIDQONT 100
Db	341	QQLSSNYYINNEDLNRN---MTVYVNPQKTYTKETFTVNLGTGYKFNPDAK----- 389
Qy	101	SIKYKYVDNAADLSESYFVNPEPEDVTNSVNIITFPNPQ-YKVEFNTPDQITTPYIV 159
Db	390	NFKIYEVDNQMFVDSFTPDTSKLDVGTQPDVITYSNDKNKTATVDLLNGQSSDKQYIIQ 449
Qy	160	VNGHIDPNSKGDALRLSTLYGCYNSNIIWRSMWDNEVAFNNGSGSGDIDK 210
Db	450	OVAYXPDNSSDTNGKIDYTLTONG-----KSSWNSYSNVNGSSSTANGDOK 495

RESULT 14

T30856 protein F2 - Streptococcus pyogenes

C:Species: Streptococcus pyogenes

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C:Accession: T30856

R:Jaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanski, E. Mol. Microbiol. 21, 373-384, 1996

A:Title: Protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes, p

A:Reference number: Z20907; MUID:97011581; PMID:8958591

A:Accession: T30856

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1039 <JAF>

A:Cross-references: UNIPROT:P72534; EMBL:U31980; NID:g1654115; PID:g1654116; PIDN:AAC445

C:Genetics:

A:Note: PRTE2

Query Match 9.5%; Score 110; DB 2; Length 1039;
Best Local Similarity 21.3%; Pred. No. 3.6;
Matches 49: Conservative 43; Mismatches 98; Indels 40; Gaps 9;

Qy	8	IDPENVKKTGNVTIATIGSTTANKTVLVDYKKGKFNLSIKGTHIDOIDKTNNTYRQTI	67
Dd	167	ICGGIEITYKGTWNVLYGNESTKESNYIINGLSNVG-----GSISYNTETGEFVWV	218
Qy	68	YNVPSGDNVIAPLYLTGNL-----KPTDSNALIDQONTSIKYKVYDNDADLESFYV	119
Dd	219	YVNPRTNIPYAVL--NLWGFAKTAQGENSNLVSSAQLTGYYDIYEYPHNYRPUTSYGV	276
Qy	120	NPNPEFDVTSNVIITFPNEN-----QYKVEFNTDDQITTPYIVVNGHIDPNSKGDLA	173
Dd	277	DISRL-NLRKDLEAKPGOSTQGANKRLRIDFG--ENLGKAFVVKVTKGDAQSQCK-ELI	332
Qy	174	LRLSTLYGYSNIIMRS-----MSWDNEVAFNNGSGSGDGIDKPVPV	214
Dd	333	VOSHLSSFN--WCQSKYLRENHSVSTNETALSPSKGSGSGSEFTKPF	379

RESULT 15

AC1533
surface protein (LPXTG motif) [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AC1533
R:Glaser, P.; Frangeul, L.; Buchrteser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schluster, T.; Simoes, N.; Tisseret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative Genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1533
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1386 <GUA>
A:Cross-references: UNIPROT:Q92DL0; GB:AI592022; PIN:CAC96035.1; PID:g16413254; GSPDB:GN
A:Experimental source: strain Clip11262
C:Genetics:
C:Gene: lin0803

Query Match 9.5%; Score 109.5; DB 2; Length 1386;
Best Local Similarity 21.0%; Pred. No. 5.8;
Matches 57: Conservative 29; Mismatches 68; Indels 117; Gaps 10;

Qy	8	IDENVKKTGNVLIAT-----GIGSTTRANKTVLVDY-----	38
Dd	1066	IDENSVKVGSDVIQTGPKYRITYSIBG-ASTTITVTVLADQSNLVAKNSTIYVGDWKQS	1124
Qy	39	-----EKYGFPYNLS--IKGTIDOIDKNTNTRYQTIVN-----	70
Dd	1125	KDNFVSATDKYGPEDISLLTIVGTVD--TTTPGEYITYSVNGLTITIIVVKENQANI	1182
Qy	71	-----PSGDNVIAPV-----LTGNLKPNTDGN-----	92
Dd	1183	VAEDSTTHAKESMKAADFVSATDKTGKINGLSSVSTGVKDINTPGNYEITYTIDGVST	1242
Qy	93	-----ALIDOONTSIKV---KYVDNAADLSESYFNVPENFEVDVNSVNITTFNPNN	139
Dd	1243	TITVTVLNHQSIOEAHDAKIKVGTSWPDEPSFLAKKFGETAFSDVTVEGTVDITYXPG	1302
Qy	140	QYKVFEFTPDQITPTPYIVVVVNGHIDNSKG	170
Dd	1303	KYOITTYIDGVSVTITIV-----EDNSKG	1327

Search completed: November 3, 2004, 19:04:07
Job time : 8.41939 secs

This Page Blank (uspto)